Accepted Manuscript

Pushing the Frontier of Minimality

Guillaume Escamocher, Barry O'Sullivan

PII: S0304-3975(18)30416-X
DOI: https://doi.org/10.1016/j.tcs.2018.06.008
Reference: TCS 11631

To appear in: *Theoretical Computer Science*

Received date: 20 June 2017
Revised date: 5 December 2017
Accepted date: 4 June 2018

Please cite this article in press as: G. Escamocher, B. O'Sullivan, Pushing the Frontier of Minimality, *Theoret. Comput. Sci.* (2018), https://doi.org/10.1016/j.tcs.2018.06.008

Pushing the Frontier of Minimality

Guillaume Escamocher*, Barry O'Sullivan*

Insight Centre for Data Analytics, University College Cork, Ireland

Abstract

The Minimal Constraint Satisfaction Problem, or Minimal CSP for short, arises in a number of real-world applications, most notably in constraint-based product configuration. It is composed of the set of CSP problems where every allowed tuple can be extended to a solution. Despite the very restrictive structure, computing a solution to a Minimal CSP instance is NP-hard in the general case. In this paper, we look at three independent ways to add further restrictions to the problem. First, we bound the size of the domains. Second, we define the arity as a function on the number of variables. Finally we study the complexity of computing a solution to a Minimal CSP instance when not just every allowed tuple, but every partial solution smaller than a given size, can be extended to a solution. In all three cases, we show that finding a solution remains NP-hard. All these results reveal that the hardness of minimality is very robust.

Keywords: constraint satisfaction, minimal CSP, NP-hardness result

1. Introduction

An instance of the Minimal Constraint Satisfaction Problem, or Minimal CSP for short, is a CSP instance where each tuple allowed in a constraint relation is part of at least one solution [9]. Since all Minimal CSP instances are satisfiable, solving such an instance does not refer to the decision problem of determining whether it has a solution, but to the exemplification of a solution.

Minimal CSP is often found 'naturally' in configuration problems [8]. A seller might want to offer its customers a large degree of customization. If, for example, the product sold is a car, some possible options might be the color of the vehicle and whether it is automatic or manual. If after choosing "automatic", "red" remains a valid option for the color parameter, then it is preferable that at least one red automatic car can be configured. The Minimal CSP can answer a number of queries relevant to product configuration in polynomial time [6], such as whether a solution exists that satisfies a given unary constraint, or whether an assignment to k variables is consistent in a Minimal CSP where all constraints are defined over k-tuples of the variables. These queries can be answered simply by inspecting the constraints of the problem instance. However, answering queries over arbitrary assignments to the variables

*Corresponding author

Email addresses: guillaume.escamocher@insight-centre.org (Guillaume Escamocher), barry.osullivan@insight-centre.org (Barry O'Sullivan)

remains hard, which has given rise to many studies of the use of automata and decision diagrams to reason about the solution sets of complex configuration problems [2].

The notion of minimality is related to that of robustness [1, 7]. Robust CSP is the problem of determining whether every partial solution of a given size can be extended to a full solution, in effect checking the minimality of an instance. On the other hand, Minimal CSP already assumes that this condition is fulfilled and instead requires to find a solution.

The restrictions defining minimality can be viewed as extreme forms of consistency. The concept of minimality is that all values not belonging to a solution have been pruned, all constraints allowing values that cannot possibly appear in a same solution have been adjusted. Yet, even though minimality offers an abundance of data about the CSP instances it applies to, it turns out that algorithms cannot use this information to significantly distinguish Minimal CSP instances from general CSP ones. Indeed, we prove that when bounding the size of the domains by a constant d and the arity of the constraints by a constant k, the Minimal CSP and the general CSP are NP-hard for the exact same values of d and k.

We also expand the concept of minimality, to study if hardness is conserved. We present the different directions that we considered. Our main result is the one revolving around what seems like the most natural expansion. In our new class of Minimal CSP instances, we significantly increase the number of sets of compatible values that can be extended to a solution. While one may think this leads to triviality, or at least tractability, we show that again no algorithm can exploit this new information in a useful way, unless P=NP. The long-term objective of this work is to identify the frontier of intractability for Minimal CSP.

Each of the three next sections of the paper presents a particular way to further restrict the Minimal CSP. In Section 2, we start by formally defining both the general CSP and the Minimal CSP, then proceed to study the complexity of the Minimal CSP when bounding the arity of the constraints and the size of the domains. In particular, we present a complexity classification over these two parameters that extends Gottlob's complexity result [6] to instances with very small domain sizes. The contents of this section have been previously published [5]. In Section 3 we provide a look into the behavior of the Minimal CSP with global constraints. Section 4 focuses on generalizing the core notion of extendable tuple in the definition of Minimality to extendable partial solution. We begin in Section 4.1 by formalizing and illustrating the new concepts that we introduce. Then in Section 4.2 we present the main result of the paper, showing that the inherent hardness of minimality is conserved even with considerable additional restrictions. Finally, we conclude in Section 5 by summarizing our contributions and outlining some future work in this area.

2. Bounding the Size of the Domains and the Arity

The first of our three generalizations of minimality deals with the arity and size parameters. Before presenting our complexity proofs, we start by formally defining the core notions of the Constraint Satisfaction Problem. In particular, we highlight the fact that we do not view the constraints of a CSP instance as a list of forbidden tuples, as is standard in the constraint literature, but instead as the complete specification of the value of every tuple of size smaller or equal than the arity, where the value here means either allowed or forbidden. Our main reason for doing so is to emphasize the role of allowed tuples, which are central to the notion of minimality but mostly ignored by the conventional CSP definition.

2.1. Definitions

We recall the definition of the Constraint Satisfaction Problem, or CSP.

Definition 1 (CSP). A *CSP instance* I comprises:

1. A set $V = \{v_1, \ldots, v_n\}$ of n variables.
2. A set $A = \{A_{v_1}, \ldots, A_{v_n}\}$ of n domains. For all $i \in [1, n]$, $A_{v_i} = \{a_1, \ldots, a_{d_i}\}$ contains the d_i possible values for the variable v_i.
3. An integer k and a set $C = \{C_1, \ldots, C_m\}$ of m constraints. To each constraint C_i is associated a different *scope* $W_i = \{w_1, \ldots, w_{k_i}\} \subseteq V$,with $2 \leq k_i \leq k$, and a set U_i of k_i-tuples from $A_{w_1} \times A_{w_2} \times \cdots \times A_{w_{k_i}}$. We say that these tuples are *allowed*, that the tuples from $A_{w_1} \times A_{w_2} \times \cdots \times A_{w_{k_i}}$ that are not in U_i are *forbidden* and that k_i is the *arity* of the constraint C_i.
   For each set $V' \subseteq V$ containing k' variables with $2 \leq k' \leq k$, there is exactly one constraint in C whose scope is exactly V' (so $m = \sum_{k'=2}^{k} \binom{n}{k'}$). We say that k is the arity of the instance.

Note that since the scopes of the constraints cover all possible sets of variables of size between 2 and the arity of the instance, defining the constraints for a given k-ary CSP instance I is equivalent to specifying whether each tuple of k' values, with $2 \leq k' \leq k$, is allowed or forbidden.

Throughout the paper, and as long as the context is clear, we will associate a tuple of assignments with the corresponding set of values. For example, we will associate the 3-tuple composed of the value a_1 assigned to the variable v_1, the value a_2 assigned to the variable v_2 and the value a_3 assigned to the variable v_3 with the set $B = \{a_1, a_2, a_3\}$, as long as it is clear that the value $a_i \in B$ is the same as the value $a_i \in A_{v_i}$, for $i = 1, 2, 3$.

A *compatible* tuple B, or *partial solution*, is a set of assignments to variables of I such that no subset of B is a forbidden tuple. Similarly, an *incompatible* tuple B is a set of assignments to variables of I such that there is a subset of B which is a forbidden tuple. We also say that values in a compatible (respectively incompatible) tuple are compatible (respectively incompatible) with each other. In particular, any value in a forbidden tuple B is incompatible with the other values in B.

A *solution*, or *full solution*, to I is a partial solution on V. We now formally define the Minimal Constraint Satisfaction Problem, or Minimal CSP.

Definition 2 (Minimal CSP). A CSP instance $I = \{V, A, C\}$ is a *Minimal CSP instance* if and only if: $\forall C_i \in C$, $\forall u$ such that u is an allowed tuple on the scope of C_i, there is at least one solution S to I such that S contains u.

In this paper, we only consider Minimal CSP instances with non-empty domains and such that each value in each domain belongs to at least one allowed tuple.

2.2. Complexity Results

Not only are Minimal CSP instances always satisfiable, but they typically contain many solutions. A Minimal CSP instance will contain as least as many solutions as there are allowed k-tuples in its least constrained constraint. For this reason, Minimal CSP instances will often be very trivial to solve. Yet, computing a solution to a Minimal CSP instance is

NP-hard [6].[1] The proof given by Gottlob [6] is a reduction from 3-SAT to a set of CSP instances M_9 such that for each instance $I \in M_9$:

- I is either a Minimal CSP instance or unsatisfiable.
- I contains at most 9 values in each domain.

This is stronger than just NP-hardness, the actual result is that computing a solution to a Minimal CSP instance is NP-hard, even when bounding the size of the domains by a fixed integer $d \geq 9$. We now further generalize this result to any $d \geq 3$:

Theorem 1. *Computing a solution to a Minimal CSP instance is NP-hard, even when bounding the size of the domains by a fixed integer $d \geq 3$.*

Proof. The proof uses the same idea as the standard NP-Completeness proof for 3SAT which transforms a clause $C = l_1 \vee l_2 \vee l_3 \vee l_4$ into two smaller clauses $C_1 = l_1 \vee l_2 \vee x$ and $C_2 = \overline{x} \vee l_3 \vee l_4$.

Suppose that we have a k-ary Minimal CSP instance I. Let v be a variable in I such that the domain of v is of size $d_v > 3$. We replace v by two new variables v_1 and v_2 to obtain a new k-ary instance I'. In the following definition, the first point defines the domains of I', while points 2. to 5. specify which k'-tuples are allowed and which are forbidden, for each k' such that $2 \leq k' \leq k$ (therefore defining the constraints of I'). As mentioned before, we will use a set notation to represent tuples in this proof and some others, since the context is clear and the order of assignments within a tuple does not matter.

1. With $\{a_1, a_2, \ldots, a_{d_v}\}$ being the domain of v, we set the domain of v_1 to $\{a_1, a_2, x\}$ and the domain of v_2 to $\{\overline{x}, a_3, \ldots, a_{d_v}\}$. The domains of the other variables remain unchanged.
2. Let $B = \{b_1, b_2, b_3, \ldots, b_{k'}\}$ be a k'-tuple such that b_1 is in the domain of v_1 and b_2 is in the domain of v_2. Then B is allowed if and only if: $b_1 = x$, $b_2 \neq \overline{x}$ and there exists some value b' such that the k'-tuple $B' = \{b', b_2, b_3, \ldots, b_{k'}\}$ is allowed in I; or $b_1 \neq x$, $b_2 = \overline{x}$ and there exists some value b' such that the k'-tuple $B' = \{b_1, b', b_3, \ldots, b_{k'}\}$ is allowed in I.
3. Let b be a value not in the domain of v_1 or v_2. Let $B = \{b, b', b_3, \ldots, b_{k'}\}$ be a k'-tuple such that b' is in the domain of v_1 and B does not contain any value from the domain of v_2. If $b' = a_i$ for $1 \leq i \leq 2$, then B is allowed in I' if and only if B was allowed in I. If $b' = x$, let $B_i = \{b, a_i, b_3, \ldots, b_{k'}\}$ for $3 \leq i \leq d_v$. B is allowed in I' if and only if one of the B_i was allowed in I.
4. Let b be a value not in the domain of v_1 or v_2. Let $B = \{b, b', b_3, \ldots, b_{k'}\}$ be a k'-tuple such that b' is in the domain of v_2 and B does not contain any value from the domain of v_1. If $b' = a_i$ for $3 \leq i \leq d_v$, then B is allowed in I' if and only if B was allowed in I. If $b' = x$, let $B_i = \{b, a_i, b_3, \ldots, b_{k'}\}$ for $1 \leq i \leq 2$. B is allowed in I' if and only if one of the B_i was allowed in I.

[1] We are aware that finding a solution to a Minimal CSP is both a *search* problem (find a solution) and a *promise* problem (the input CSP is satisfiable because it is minimal), rather that a *decision* problem. However, we use this terminology in the same manner as Gottlob [6], where a thorough discussion of the matter can be found.

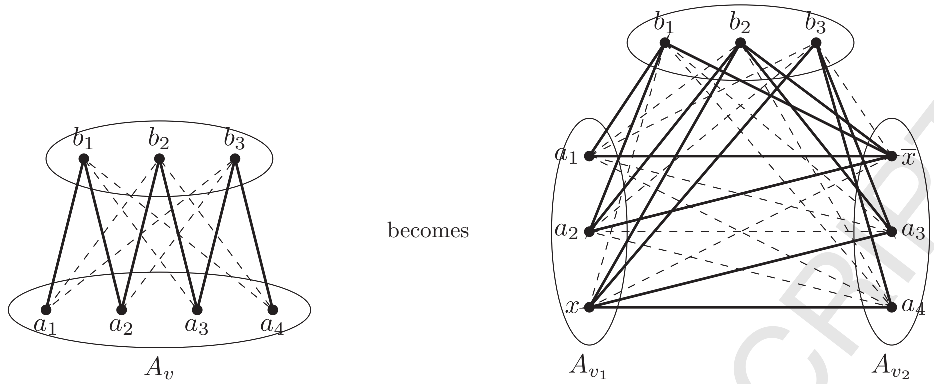

Figure 1: Transforming a variable v into two variables v_1 and v_2 with smaller domains.

5. All other k'-tuples in I' remain as they were in I.

Figure 1 illustrates an example of the transformation for $k = 2$ and $d_v = 4$. A_v, A_{v_1} and A_{v_2} denote the domains of v, v_1 and v_2 respectively. The continuous lines represent allowed pairs, while the dashed lines represent forbidden pairs.

When reducing I to I', we remove one variable and add two. So the number of variables in I' is $n + 1$, with n being the number of variables in I. We also add two new values, x in the domain of v_1 and $\overline{x}$ in the domain of v_2. So the number of values in I' is $d_t + 2$, with d_t being the total number of values in I. The arity of I' is the same as the arity of I. So the number of tuples in the constraints of I' is polynomial in the number of tuples in the constraints of I. Furthermore, specifying whether each tuple of I' is allowed or forbidden only requires us to check some corresponding tuples from I. Therefore, reducing I to I' is done in polynomial time in the size of I. So in order to obtain the desired result, it only remains to prove that I' is a Minimal CSP instance and that finding a solution for I' allows us to find in polynomial time a solution for I.

1. I' is a Minimal CSP instance:
   We have to prove that each allowed k'-tuple in I' is in a solution for I', for each k' such that $2 \leq k' \leq k$. Let k' be an integer such that $2 \leq k' \leq k$ and let $B = \{b_1, b_2, b_3, \ldots, b_{k'}\}$ be an allowed k'-tuple in I' such that:
   - Either b_1 is in the domain A_{v_1} of v_1, or B does not contain any value from A_{v_1}.
   - Either b_2 is in the domain A_{v_2} of v_2, or B does not contain any value from A_{v_2}.

   We necessarily have one of the following seven cases:
   (a) $b_1 = x$ and $b_2 = a_i$ for some $i \in [3, \ldots, d_v]$: from the second point in the definition of I', we know that there is some allowed k'-tuple $B' = \{b', a_i, b_3, \ldots, b_{k'}\}$ in I. Since I is minimal, there is some solution S for I such that B' belongs to S. Let $S' = S \cup \{x\}$. By construction, S' is a solution for I' containing B.
   (b) $b_1 = a_i$ for some $i \in [1, 2]$ and $b_2 = \overline{x}$: same argument as for (a).
   (c) $b_1 = x$ and $b_2 \notin A_{v_2}$: from the third point in the definition of I', we know that there is some allowed k'-tuple $B_i = \{a_i, b_2, b_3, \ldots, b_{k'}\}$ in I, with $i \in [3, \ldots, d_v]$. Since I is minimal, there is some solution S for I such that B_i belongs to S. Let $S' = S \cup \{x\}$. By construction, S' is a solution for I' containing B.

(d) $b_1 \notin A_{v_1}$ and $b_2 = \overline{x}$: same argument as for (c), using the fourth point in the definition of I' instead of the third.

(e) $b_1 \in A_{v_1}$, $b_1 \neq x$ and $b_2 \notin A_{v_2}$: from the third point in the definition of I', we know that B is allowed in I too. Since I is a Minimal CSP instance, there is some solution S for I such that B belongs to S. Let $S' = S \cup \{x\}$. By construction, S' is a solution for I' containing B.

(f) $b_1 \notin A_{v_1}$, $b_2 \in A_{v_2}$ and $b_2 \neq \overline{x}$: same argument as for (e), using the fourth point in the definition of I' instead of the third.

(g) $b_1 \notin A_{v_1}$ and $b_2 \notin A_{v_2}$: from the fifth point in the definition of I', we know that B is allowed in I too. Since I is a Minimal CSP instance, there is some solution S for I such that B belongs to S. Let a be the point of S in A_v. If $a = a_1$ or $a = a_2$, then let $S' = S \cup \{\overline{x}\}$. Otherwise, let $S' = S \cup \{x\}$. By construction, S' is a solution for I' containing B.

So if B is an allowed k'-tuple of I', then B is in a solution for I'.

2. Finding a solution for I from a solution for I':
   Suppose that we have a solution S' for I'. From the second point in the definition of I', we know that there is no allowed k-tuple containing both x and $\overline{x}$. Therefore, S' must contain one of the a_i. Let $S = S' \backslash \{x\}$ (or $S = S' \backslash \{\overline{x}\}$ if $\overline{x} \in S'$). By construction, S is a solution for I.

The domain size of v_1 is $3 < d_v$, and the domain size of v_2 is $d_v - 1 < d_v$. Therefore, if we have a variable with a domain of size strictly greater than 3, then we can replace it by two variables with strictly smaller domain size. By iteratively applying this operation until all domains have a size of 3 or less, we can reduce I to a Minimal CSP instance where the size of all domains is bounded by 3. Since computing a solution to a Minimal CSP instance is NP-hard [6], we have the result. □

In the case of Boolean domains, we also have NP-hardness if the arity of the instances is $k \geq 3$.

Theorem 2. *For all $k > 2$, computing a solution to a k-ary Minimal CSP instance is NP-hard, even when bounding the size of the domains by 2.*

Proof. The proof is based on the transformation from a domain $A = \{a_1, a_2, a_3\}$ of size 3 to three domains $A_1 = \{a_1, \overline{a_1}\}$, $A_2 = \{a_2, \overline{a_2}\}$, $A_3 = \{a_3, \overline{a_3}\}$, each of size 2.

Let $k > 2$. From Theorem 1, we know that computing a solution to a k-ary Minimal CSP instance is NP-hard, even when bounding the size of the domains by 3. Therefore, we just need to reduce the k-ary Minimal CSP with domain size bounded by 3 to the k-ary Minimal CSP with domain size bounded by 2. Let I be a k-ary Minimal CSP instance such that the size of each domain in I is bounded by 3. Let v be a variable in I such that the domain of v is of size 3. We replace v by three new variables v_1, v_2 and v_3 to obtain a new k-ary instance I'. In the following definition, the first point defines the domains of I', while points 2. to 7. specify which k'-tuples are allowed and which are forbidden, for each k' such that $2 \leq k' \leq k$ (therefore defining the constraints of I').

1. With $\{a_1, a_2, a_3\}$ being the domain of v, we set the domain of v_i to be $A_{v_i} = \{a_i, \overline{a_i}\}$ for $1 \leq i \leq 3$. The domains of the other variables remain unchanged.
2. All k'-tuples in I' containing both a_i and a_j for some $1 \leq i \neq j \leq 3$ are forbidden.

3. Let $B = \{\overline{a_i}, \overline{a_j}, b_1, b_2, \ldots, b_{k'-2}\}$ be a k'-tuple for some $1 \leq i \neq j \leq 3$. Let h be the integer between 1 and 3 such that $h \neq i$ and $h \neq j$. B is allowed if and only if there is some allowed k'-tuple in I containing a_h and all the b_g for $1 \leq g \leq k'-2$. Note that this covers the particular cases when a_h is one of the b_g (in which case B is allowed if and only if there is an allowed k'-tuple in I containing all the b_g) and when $\overline{a_h}$ is one of the b_g (in which case B is forbidden because $\overline{a_h}$ does not appear in I).
4. Let $B = \{a_i, b_1, b_2, \ldots, b_{k'-1}\}$ be a k'-tuple such that $1 \leq i \leq 3$ and no b_j is in the domain of v_1, v_2 or v_3. B is allowed in I' if and only if B is allowed in I.
5. Let $B = \{\overline{a_i}, b_1, b_2, \ldots, b_{k'-1}\}$ be a k'-tuple such that $1 \leq i \leq 3$ and no b_j is in the domain of v_1, v_2 or v_3. B is allowed if and only if there is some $j \neq i$, with $1 \leq j \leq 3$, such that $\{a_j, b_1, b_2, \ldots, b_{k'-1}\}$ is allowed in I.
6. Let $B = \{a_i, \overline{a_j}, b_1, b_2, \ldots, b_{k'-2}\}$ be a k'-tuple such that $1 \leq i \neq j \leq 3$ and no b_h is in the domain of v_1, v_2, or v_3. B is allowed if and only if there is some $b_{k'-1}$ such that $\{a_i, b_{k'-1}, b_1, b_2, \ldots, b_{k'-2}\}$ is allowed in I.
7. All other k'-tuples in I' remain as they were in I.

When reducing I to I', we remove one variable and add three. So the number of variables in I' is $n+2$, with n being the number of variables in I. We also add three new values, $\overline{a_1}$, $\overline{a_2}$ and $\overline{a_3}$ in the domains of v_1, v_2 and v_3 respectively. So the number of values in I' is $d_t + 3$, with d_t being the total number of values in I. The arity of I' is the same as the arity of I. So the number of tuples in the constraints of I' is polynomial in the number of tuples in the constraints of I. Furthermore, specifying whether each tuple of I' is allowed or forbidden only requires us to check some corresponding tuples from I. Therefore, reducing I to I' is done in polynomial time in the size of I. So in order to obtain the desired result, it only remains to prove that I' is a Minimal CSP instance and that finding a solution for I' allows us to find in polynomial time a solution for I.

1. I' is a Minimal CSP instance:
   We have to prove that each allowed k'-tuple in I' is in a solution for I', for each k' such that $2 \leq k' \leq k$. Let k' be an integer such that $2 \leq k' \leq k$ and let B be an allowed k'-tuple in I'. We necesarily have one of the six following cases:
   (a) $B = \{a_i, b_1, b_2, \ldots, b_{k'-1}\}$ with $1 \leq i \leq 3$ and neither b_j in the domain of v_1, v_2 or v_3. Without loss of generality, we assume that $i = 1$. From the fourth point in the definition of I', we know that B is also allowed in I. Since I is a Minimal CSP instance, there is some solution S for I such that B belongs to S. Let $S' = S \cup \{\overline{a_2}, \overline{a_3}\}$. By construction, S' is a solution for I' containing B.
   (b) $B = \{a_i, \overline{a_j}, b_1, b_2, \ldots, b_{k'-2}\}$ with $1 \leq i \neq j \leq 3$ and no b_h in the domain of v_1, v_2 or v_3. Without loss of generality, we assume that $i = 1$. From the sixth point in the definition of I', we know that there is an allowed k'-tuple in I containing a_1 and all b_h for $1 \leq h \leq k'-1$. Since I is a Minimal CSP instance, there is some solution S for I such that both a_1 and all the b_h belong to S. Let $S' = S \cup \{\overline{a_2}, \overline{a_3}\}$. By construction, S' is a solution for I' containing B.
   (c) $B = \{a_i, \overline{a_j}, \overline{a_h}, b_1, b_2, \ldots, b_{k'-3}\}$ with $1 \leq i, j, h \leq 3$ and i, j, h all distinct. Without loss of generality, we assume that $i = 1$. From the third point in the definition of I', we know that there is an allowed k'-tuple in I containing a_1 and all the b_g for $1 \leq g \leq k-3$. Since I is a Minimal CSP instance, there is a solution S for I such that a_i and all the b_g belong to S. Let $S' = S \cup \{\overline{a_2}, \overline{a_3}\}$. By construction, S' is a solution for I' containing B.

(d) $B = \{\overline{a_i}, b_1, b_2, \ldots, b_{k'-1}\}$ with $1 \leq i \leq 3$ and no b_j in the domain of v_1, v_2 or v_3. Without loss of generality, we assume that $i = 1$. From the fifth point in the definition of I', either $\{a_2, b_1, b_2, \ldots, b_{k'-1}\}$ or $\{a_3, b_1, b_2, \ldots, b_{k'-1}\}$ is allowed in I. Without loss of generality, we assume the former. Since I is a Minimal CSP instance, there is some solution S for I such that $\{a_2, b_1, b_2, dots, b_{k'-1}\}$ belongs to S. Let $S' = S \cup \{\overline{a_1}, \overline{a_3}\}$. By construction, S' is a solution for I' containing B.

(e) $B = \{\overline{a_i}, \overline{a_j}, b_1, b_2, \ldots, b_{k'-2}\}$ with $1 \leq i \neq j \leq 3$ and no b_h not in the domain of v_1, v_2 or v_3. Without loss of generality, we assume that $i = 1$ and $j = 2$. From the third point in the definition of I', there is an allowed k'-tuple in I containing a_3 and all the b_h for $1 \leq h \leq k-2$. Since I is a Minimal CSP instance, there is some solution S for I such that a_3 and all the b_h belong to S. Let $S' = S \cup \{\overline{a_1}, \overline{a_2}\}$. By construction, S' is a solution for I' containing B.

(f) $B = \{b_1, b_2, \ldots, b_{k'}\}$ with no b_i being in the domain of v_j, for any $1 \leq i, j \leq 3$. From the seventh point in the definition of I', we know that B is allowed in I too. Since I is a Minimal CSP instance, there is some solution S for I such that B belongs to S. Let a_i be the point of S in the domain of v. Without loss of generality, we assume that $i = 1$. Let $S' = S \cup \{\overline{a_2}, \overline{a_3}\}$. By construction, S' is a solution for I' containing B.

So if B is an allowed k'-tuple of I', then B is in a solution for I'.

2. Finding a solution for I from a solution for I':
Suppose that we have a solution S' for I'. From the second and third points in the definition of I', we know that S' must contain a_i, $\overline{a_j}$ and $\overline{a_h}$, for some distinct i, j and h between 1 and 3. Without loss of generality, we assume that $i = 1$. Let $S = S' \backslash \{\overline{a_2}, \overline{a_3}\}$. By construction, S is a solution for I.

Therefore, if we have a variable with a domain of size 3, then we can replace it by three variables with domains of size 2. By iteratively applying this operation until all domains have a size of 2 or less, we can reduce I to a Minimal CSP instance where the size of all domains is bounded by 2. So we have the result.

□

The binary Boolean Minimal CSP is polynomial, since the more general binary Boolean CSP can be trivially reduced to 2-SAT, which is polynomial [12]. Combined with Theorems 1 and 2, and with the triviality of CSP instances consisting entirely of single-valued variables, we can complete the classification:

Theorem 3. *Computing a solution to a k-ary Minimal CSP instance when the size of the domains is bounded by d is NP-hard if and only if ($d \geq 3$ or ($d = 2$ and $k \geq 3$)).*

The result is summarized in Table 1.

Corollary 1. *When bounding the size of the domains by a constant d and the arity of the constraints by a constant k, Minimal CSP and the general CSP are NP-hard for the exact same values of d and k.*

3. Relativistic Arity

We have considered scenarios in which we consider CSPs involving constraints of constant arity. However, a powerful aspect of constraint programming [11] is due to the use of global constraints [13]. Informally, a global constraint defines a relation between a set of variables of non-fixed arity. The classic example is the `alldifferent`$(x_1, \ldots, x_k)$ constraint [10], which states the variables $x_1, \ldots, x_k$ must take all different values. While this is equivalent to stating that these variables must be pairwise different, a relation that can be specified as a clique of binary inequality constraints, we can take advantage of the more global view that the `alldifferent` constraint gives us. Specifically, we are interested in pruning from the domains of each x_i any value that does not participate in at least one solution of the `alldifferent`. When the domains of the variables contain only values that participate in a solution to the constraint, we say that we have achieved *generalised arc consistency* (GAC) [4]. For many global constraints GAC can be achieved in polynomial time, while for other constraints it is NP-hard.

We now turn our focus on minimality involving constraints for which the arity is not fixed, but rather is a result of a function applied to the number of variables in the instance.

Definition 3 (α-arity). Let α be a function from $\mathbb{N}$ to $\mathbb{N}$. Let I be a CSP instance with n variables. We say that I is an *α-ary* CSP instance if the arity of I is $(\alpha(n))$.

Definition 4 (α-ary Minimal CSP). Let α be a function from $\mathbb{N}$ to $\mathbb{N}$. Let I be an α-ary CSP instance with n variables. We say that I belongs to the *α-ary Minimal Constraint Satisfaction Problem* (or α-ary Minimal CSP for short) if every allowed set of $\alpha(n)$ values of I can be extended to a solution for I.

After having formally defined this new form of minimality, we now present a complexity result for such a class of Minimal CSP instances. Note that this class neither contains nor is contained in any class that we previously studied.

Theorem 4. *Let $p > 0$ be an integer. Let $\alpha : N \to \lfloor p\sqrt{N} \rfloor$ be a function from $\mathbb{N}$ to $\mathbb{N}$. Then computing a solution to an α-ary Minimal CSP instance is NP-hard. Furthermore, the result holds even when bounding the size of the domains by $d = 3$.*

Proof. We first prove the result for $p = 2$. We are going to reduce the binary Minimal CSP on domains of size 3 to the α-ary Minimal CSP on domains of size 3.

Let I be a binary Minimal CSP instance with n variables $v_1, \ldots, v_n$, and with domains of size $d = 3$. Assume without loss of generality that the three values in each domain are 0,1 and 2.

Table 1: Complexity of the k-ary Minimal CSP with domains of size bounded by d.

d \ k	2	≥ 3
1	tractable	tractable
2	tractable	NP-hard
≥ 3	NP-hard	NP-hard

Let $n' = n^2$. We then build a $(2n)$-ary CSP instance I' on n' variables $v'_1, \ldots, v'_{n'}$. We set each domain in I' to be composed of the three values 0,1 and 2. Let $A = \{a_1, a_2, \ldots, a_{2n}\}$ be a $(2n)$-tuple in I'. Let i and j be such that for each $1 \le q \le n$, a_q is in the domain of the variable $v'_{(i-1)n+q}$ and a_{n+q} is in the domain of the variable $v'_{(j-1)n+q}$. Let $a = (a_1 + a_2 + \cdots + a_n) \mod 3$ be the sum modulo 3 of the values of A that are in the domains of the variables $v'_{(i-1)n+1}, \ldots, v'_{in}$ and let $b = (a_{n+1} + a_{n+2}+, \ldots, a_{2n}) \mod 3$ be the sum modulo 3 of the values of A that are in the domains of the variables $v'_{(j-1)n+1}, \ldots, v'_{jn}$. Then we set A to be allowed in I' if and only if the value a from the domain of v_i is compatible with the value b from the domain of v_j in I. All other $(2n)$-tuples in I' are set to be allowed.

The intuitive idea behind the construction of I' is that for each variable v_i in I, there are n corresponding variables $v_{(i-1)n+1}, \ldots, v_{in}$ in I', and that the sum modulo 3 of the values assigned to these n variables in I' corresponds to the value assigned to v_i in I.

We are now going to prove that I has a solution if and only if I' has a solution. Suppose first that we have a solution $S = \{a_1, \ldots, a_n\}$ for I. Without loss of generality, assume that a_i is in the domain of v_i for each $1 \le i \le n$. Let $S' = \{a'_1, \ldots, a'_{n'}\}$ be a set of n' values of I' such that:

- For each $1 \le i \le n'$, a'_i is in the domain of v'_i.
- For each $1 \le i \le n$, $a'_{in} = a_i$.
- For each $1 \le i \le n$, the values $a'_{(i-1)n+1}, a'_{(i-1)n+2}, \ldots, a'_{(i-1)n+n-1}$ are all equal to 0.

If there is a forbidden tuple A in S', then from the definition of I' there are some i and j such that $A = \{a'_{(i-1)n+1}, a'_{(i-1)n+2}, \ldots, a'_{in}\} \cup \{a'_{(j-1)n+1},\ a'_{(j-1)n+2}, \ldots, a'_{jn}\}$ and the values $(a'_{(i-1)n+1} + a'_{(i-1)n+2} + \cdots + a'_{(i-1)n+n-1} + a'_{in}) \mod 3$ in the domain of v_i and $(a'_{(j-1)n+1} + a'_{(j-1)n+2} + \cdots + a'_{(j-1)n+n-1} + a'_{jn}) \mod 3$ in the domain of v_j are incompatible in I. So the values $(0 + 0 + \cdots + 0 + a_i) \mod 3$ in the domain of v_i and $(0 + 0 + \cdots + 0 + a_j) \mod 3$ in the domain of a_j are incompatible in I. So a_i and a_j are incompatible in I, which is in contradiction with the assumption that S is a solution for I. Therefore, there is no forbidden tuple in S'. Therefore, S' is a solution for I'. So if we have a solution for I, then we have a solution for I'.

Suppose now that we have a solution $S' = \{a'_1, \ldots, a'_{n'}\}$ for I'. Without loss of generality, assume that a'_i is in the domain of v'_i for each $1 \le i \le n'$. Let $S = \{a_1, \ldots, a_n\}$ be a set of n values of I such that:

- For all $1 \le i \le n$, a_i is in the domain of v_i.
- For all $1 \le i \le n$, a_i is equal to the sum modulo 3 of the n values $a'_{(i-1)n+1}, \ldots, a'_{in}$.

If two values a_i and a_j from S are incompatible, then from the definition of I' the tuple $\{a'_{(i-1)n+1}, a'_{(i-1)n+2}, \ldots, a'_{in}\} \cup \{a'_{(j-1)n+1}, a'_{(j-1)n+2}, \ldots, a'_{jn}\}$ is forbidden and we have a contradiction. So there is no forbidden tuple within S. So S is a solution for I. So if we have a solution for I', then we have a solution for I. So we have shown that I has a solution if and only if I' has a solution.

We are now going to prove that I' is an α-ary CSP instance. Since the arity of I' is $2n$ and $n' = n^2$, I' is an α-ary CSP instance. So in order to prove that I' is an α-ary CSP instance, we just have to prove that every allowed set of $\alpha(n')$ values of I' can be extended to a solution.

We now define several useful tools. For each $1 \leq i \leq n$, let $V_i = \{v'_{(i-1)n+1}, \ldots, v'_{in}\}$. We say that the ***scope*** of a set A of values of I' is the set W of variables of I' such that the domain of each variable of W contains at least one value of A and each value of A is in the domain of a variable of W. For each set A of n values of I' such that V_i is the scope of A for some i, let $f(A)$ be the sum modulo 3 of all the values in A.

Lemma 1. *Let A be a set of q values of I' such that $q < n$ and let W be the scope of A. Suppose that there is some i such that $W \subset V_i$. Let a be a value in the domain of v_i. Then there is a set A' such that $A \subset A'$, the scope of A' is V_i and $f(A') = a$.*

Proof. Let $A = \{a_1, a_2, \ldots, a_q\}$ be a set of a values of I' such that $q < n$. Let W be the scope of A. Suppose that there is some i such that $W \subset V_i$. Without loss of generality, assume that $W = \{v'_{(i-1)n+1}, v_{(i-1)n+2}, dots, v'_{(i-1)n+q}\}$ and that $a_{q'}$ is in the domain of $v'_{(i-1)n+q'}$ for each $1 \leq q' \leq q$. Let $a_0 = (a_1 + a_2 + \cdots + a_q) \mod 3$ be the sum modulo 3 of all values in A if A is not empty and let $a_0 = 0$ otherwise. Let a be a value in the domain of v_i.
Let $A' = \{a'_1, a'_2, \ldots, a'_n\}$ be the set of n values of I' such that :

1. For each $1 \leq r \leq n$, a'_r is in the domain of $v'_{(i-1)n+r}$.
2. For each $1 \leq r \leq q$, $a'_r = a_r$.
3. For each $q < r < n$, a'_r is equal to 0.
4. a'_n is equal to $(a - a_0) \mod 3$.

From 2., we know that $A \subset A'$. From 1., we know that the scope of A' is V_i. From 3. and 4., we know that $f(A') = a_0 + ((a - a_0) \mod 3) = a$. Therefore, we have constructed a set A' that satisfies the conditions of the lemma. □

Suppose that we have a compatible $(2n)$-tuple $A = \{a'_1, a'_2, \ldots, a'_{2n}\}$ in I'. Let W be the scope of A. Since there are only $2n$ values in A, only two sets among $V_1, V_2, \ldots, V_n$ can be fully contained in W. Therefore, there are three possibilities for the composition of W:

1. There are some i and j such that $W = V_i \cup V_j$. Let A_i be the restriction of A to V_i and let A_j be the restriction of A to V_j. Let $a_i = f(A_i)$ and let $a_j = f(A_j)$. Since A is compatible, the value a in the domain of v_i is compatible with the value b in the domain of v_j in I. Since I is minimal, this allowed tuple can be extended to a solution $S = \{a_1, a_2, \ldots, a_n\}$ for I, such that each a_q is in the domain of v_q. Let $B = \{b_1, b_2, \ldots, b_{n'}\}$ be the set of n' values of I' such that:
   (a) For each q such that $1 \leq q \leq n'$, b_q is in the domain of v'_q.
   (b) For each q such that $(i-1)n + 1 \leq q \leq in$ or $(j-1)n + 1 \leq q \leq jn$, b_q is equal to the value of A in the domain of v'_q.
   (c) For each a such that $1 \leq q \leq n$, $q \neq i$ and $q \neq j$, b_{qn} is equal to a_q.
   (d) For each q such that the value of b_q has not been specified in (b) or (c), the value of b_q is equal to 0.

   From (a) and (b), B is an extension of A to the whole instance I'. From (c), (d) and the definition of I', B does not contain any forbidden tuple. Therefore, we have extended A to a solution for I'.

2. There is some i such that $W = W_1 \cup W_2 \cup \cdots \cup W_{i-1} \cup V_i \cup W_{i+1} \cup \cdots \cup W_n$, with each W_j being a (possibly empty) strict subset of V_j. Let A_i be the restriction of A to V_i and for each $1 \leq j \leq n$ such that $j \neq i$ let A_j be the restriction of A to W_j. Let $a_i = f(A_i)$. Since I is Minimal, there is a solution $S = \{a_1, a_2, \ldots, a_n\}$ for I, such that each a_j is in the domain of v_j. From Lemma 1, we know that we can extend A to a set A' of n' values of I' such that for each $1 \leq j \leq n$ we have $f(A'_j) = a_j$, with A'_j being the restriction of A' to V_j. From the definition of I', A' is a solution for I'.
3. There is no i such that $V_i \subset W$. So $W = W_1 \cup W_2 \cup \cdots \cup W_n$, with each W_i being a (possibly empty) strict subset of V_i. For each $1 \leq i \leq n$, let A_i be the restriction of A to W_i. Let $S = \{a_1, a_2, \ldots, a_n\}$ be a solution for I, such that each a_i is in the domain of v_i. From Lemma 1, we know that we can extend A to a set A' of n' values of I' such that for each $1 \leq i \leq n$ we have $f(A'_i) = a_i$, with A'_i being the restriction of A' to V_i. From the definition of I', A' is a solution for I'.

In all cases, A can be extended to a solution for I', which proves that I' belongs to the α-ary Minimal CSP. So we have reduced the binary Minimal CSP with a bound of 3 on the size of the domains to the α-ary CSP with a bound of 3 on the size of the domains.

Both the number of variables in I' (n^2) and the size of the domains of I' (3) are polynomial in the size of I, but the number of $(2n)$-tuples in I' is equal to $\binom{n^2}{2n} \times 2^{2n}$, which is exponential in the size of I. Therefore, it is not trivial whether the reduction is done in time polynomial in the size of I. However, at no point in the reduction do we need to individually check all tuples of I'. We only need to express the constraints of I' in the form of a function which answers in time polynomial in the size of n whether a given partial assignment for I' is compatible.
Suppose that we have a partial assignment A for I'. If there is a forbidden $(2n)$-tuple A' within A, then from the definition of I' there are some i and j such that the scope of A' is $V_i \cup V_j$. There are only $n(n-1)/2$ such pairs (i, j), so we only need to check the at most $n(n-1)/2$ subsets of A that fit the description. If we have i, j and A', where A' is a subset of A and the scope of A' is $V_i \cup V_j$, then checking whether A' is allowed can be done in polynomial time in the size of I by just summing the values of A' modulo 3 and comparing the results with the constraints of I. Therefore, the size of I', including the variables, the domains and an exact representation of the constraints, is polynomial in the size of I. Therefore the reduction is done in time polynomial in the size of I.

From Theorem 3, the binary Minimal CSP is NP-hard even with a bound of 3 on the size of the domains. Therefore we have the result for $p = 2$.

For $p = 1$, the proof is the same with the addition of $3n^2$ variables in I' that are compatible with everything in I'. The number of variables in I' is now $n' = 4n^2$, and the arity of I' becomes $2n = \sqrt{n'} = \alpha(n)$.

For $p > 2$, the proof is also the same with the only difference being that the instance I at the start of the reduction is now a p-ary Minimal instance. □

4. Relativistic Minimality

In this section we look at the complexity of the Minimal CSP when we alter the composition of the set of objects that can be extended to a solution. Specifically, we now require that all partial solutions smaller than some particular size can be extended to a full solution.

The upper bound on the size of the partial solutions will be given by a function α, where the argument taken by α is the number n of variables in an instance. Each different function α defines a different class of Minimal CSP instances.

It is natural to wonder how high $\alpha(n)$ can be while still maintaining the complexity of computing a solution to be NP-hard. For example, we know from Theorem 3 that it is NP-hard to find a solution to a binary CSP instance even when requiring that all compatible pairs of values can be extended to a solution. What happens if in addition to that all compatible tuples of size 3, 4 and 5 (case $\alpha(n) = 5$) can also be extended to a solution? This new class is much more restrictive, so it is not obvious whether NP-hardness is conserved. If it is, how many further similar restrictions can we add to the class before it becomes tractable? What if $\alpha(n)$ is equal to $\log_4(n)$, or to $\sqrt[3]{n}$, or to $n/2$?

We will prove in Section 4.2 that $\alpha(n)$ can be as high as $\sqrt[\epsilon]{n}$ for any real number $\epsilon > 2$ and computing a solution will still be NP-hard. Such functions include the previously mentioned cases of $\alpha(n) = 5$, $\alpha(n) = \log_4(n)$ and $\alpha(n) = \sqrt[3]{n}$. However, the complexity for the functions $\alpha(n) = n/p$ is still open.

The notion of large partial solutions that can be extended to a full solution has been studied before in the context of robustness [1]. Their paper looked at the class of CSP instances where every compatible tuple of size smaller than some r can be extended to a full solution. However, while their bound r is allowed to be greater than the arity of the instance, it is still a constant. On the other hand, our upper bound α is actually a function on the number of variables in the instance.

4.1. Definitions

Definition 5 (p Minimality). Let $p > 0$ be an integer. We say that a k-ary CSP instance I on n variables is p *Minimal* if $k \leq p \leq n$ and every compatible p-tuple of I can be extended to a solution for I.

Definition 6 (p^- Minimality). Let $p > 0$ be an integer. We say that a k-ary CSP instance I on n variables is p^- *Minimal* if $k \leq p \leq n$ and I is a' Minimal for every integer q such that $k \leq q \leq p$.

Lemma 2. *Let I be k-ary CSP instance. Let p_1 and p_2 be two integers such that $k \leq p_1 \leq p_2$. Then:*

$$I \text{ is } p_1^- \text{ Minimal} \Leftarrow I \text{ is } p_2^- \text{ Minimal}$$

Proof. Suppose that I is p_2^- Minimal. From Definition 6, I is p Minimal for each p such that $k \leq p \leq p_2$. Since $p_1 \leq p_2$, I is p Minimal for each p such that $k \leq p \leq p_1$. So from Definition 6, I is p_1^- Minimal and we have the result. □

Definition 7 (α^- Minimality). Let I be a k-ary CSP instance on n variables. Let α be a function from $\mathbb{N}$ to $\mathbb{R}$ such that $k \leq \alpha(n) \leq n$. We say that I is α^- *Minimal* if it is $\alpha(n)^-$ Minimal.

Definition 8 (α^- Minimal CSP). Let α be a function from $\mathbb{N}$ to $\mathbb{R}$. The α^- *Minimal Constraint Satisfaction Problem* (or α^- Minimal CSP for short) is the problem of computing a solution to α^- Minimal CSP instances.

To give a simple concrete example illustrating these new notions, suppose that we want to color a map composed of the six New England states, such that at most four colors are used and two adjacent states never share the same color. Figure 2 represents the associated coloring instance. There is one variable for each state, each domain contains the four colors $\{1, 2, 3, 4\}$ and there is a *diff* constraint, represented by a dashed line, between two states if they are adjacent.

The 4-Coloring problem is always satisfiable for planar graphs [3] so this instance is satisfiable. Furthermore, any compatible assignment on two variables, that is either an assignment to two non-adjacent states or an assignment of two different colors to two adjacent states, can be extended to a full solution. Therefore, the instance is Minimal. Since all the constraints are binary, from Definition 5 the instance is 2 Minimal.

Going even further, any assignment of colors to three different states that does not violate any *diff* constraint can be extended to a solution for all six states. For example, the assignment of the color 1 to Massachusetts, the color 2 to Vermont and the color 3 to Maine forms a compatible 3-tuple that can be completed to a solution for the full instance by assigning the color 2 to Connecticut, the color 3 to Rhode Island and the color 4 to New Hampshire. Therefore, from Definition 5 the instance is 3 Minimal. Since the instance is also 2 Minimal and binary, from Definition 6 the instance is 3^- Minimal. However, there exist compatible 4-tuples that cannot be extended to a solution. For example, if we assign the color 1 to Connecticut, the color 2 to Rhode Island, the color 3 to Vermont and the color 4 to New Hampshire, then there is no assignment to Massachusetts that is compatible with all four values already picked. So from Definition 5 the instance is not 4 Minimal, and therefore from Definition 6 it is not 4^- Minimal either.

Let α_1 be the function from $\mathbb{N}$ to $\mathbb{R}$ such that $\alpha_1(n) = n/2$ for each integer n. Let α_2 be the function from $\mathbb{N}$ to $\mathbb{R}$ such that $\alpha_2(n) = 2n/3$ for each integer n. Since there are six variables in the instance and the instance is 3^- Minimal, from Definition 7 the instance is α_1^- Minimal. Since there are six variables in the instance and the instance is not 4^- Minimal, from Definition 7 the instance is not α_2^- Minimal.

4.2. Complexity Results

Our main theorem states that if α is a function and $\epsilon > 2$ some real number such that $\alpha(n) \leq \sqrt[\epsilon]{(n)}$ for all n, then the α^- Minimal CSP is NP-hard, even when the constraints are binary. We first prove the result for the special case $\epsilon = 3$.

Proposition 1. *Let γ be the function from $\mathbb{N}$ to $\mathbb{R}$ such that $\forall N \in \mathbb{N}, \gamma(N) = \sqrt[3]{N}$. Then the γ^- binary Minimal Constraint Satisfaction Problem is NP-hard.*

Proof. Let $P_{d\leq 3}$ be the binary Minimal Constraint Satisfaction Problem restricted to the instances with at most 3 values in each domain, and let $P_{d\leq 3}^{n\geq 12}$ be $P_{d\leq 3}$ restricted to the instances with at least 12 variables. From Theorem 3, we know that $P_{d\leq 3}$ is NP-hard. Since the number of instances in $P_{d\leq 3} \backslash P_{d\leq 3}^{n\geq 12}$ is finite, $P_{d\leq 3}^{n\geq 12}$ is also NP-hard. We are going to reduce $P_{d\leq 3}^{n\geq 12}$ to the γ^- binary Minimal CSP. Let I be a binary Minimal CSP instance with n variables $v_1, \ldots, v_n$, such that $n \geq 12$ and there are at most three values in the domain of v_i for each $1 \leq i \leq n$. Since adding values that are incompatible with the rest of the instance does not add allowed tuples nor remove solutions, we can add values in the domains that have strictly less than three values without losing minimality. Without loss of generality,

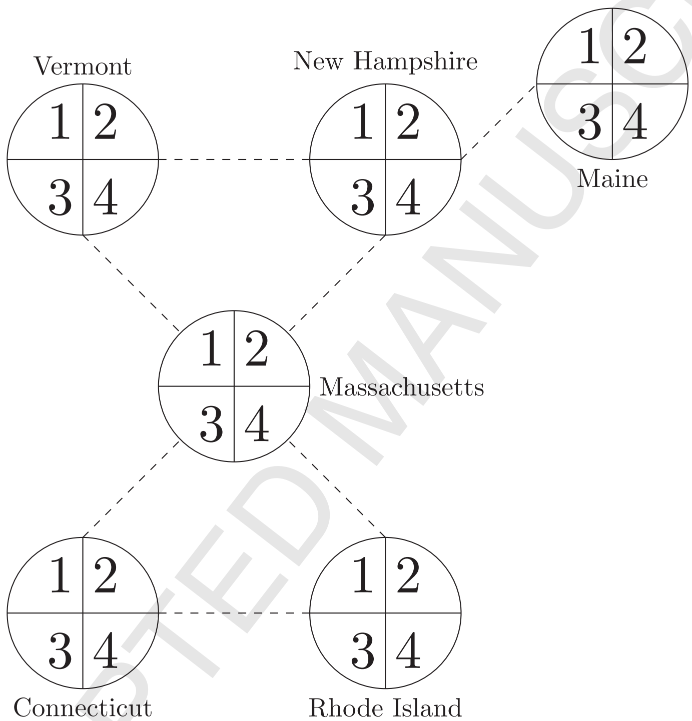

Figure 2: Adjacencies within the New England states.

assume that the domain of each variable in I is composed of the three values 0, 1 and 2. From now on, we will refer to a value from one of these domains, that is a value being equal to one of the three integers 0, 1 and 2, as a "trilean" value.

Let $\alpha = n^3$, and let $\beta = 2\alpha + 1$. Note that β is always odd. We are going to reduce I to a binary α^- Minimal CSP instance I'. The following nine bullet points describe the construction of I'. At the start of each bullet point, we give between parenthesis a few keywords describing the notion covered by this specific bullet point, so that the reader can more easily find what he/she wants when going back to the definition of I' in the future. While each bullet point is necessary for the correctness of the proof, the second bullet point is of particular importance, as it describes the main idea behind the proof.

1. (root variables) For each variable v_i in I, we create α variables $v_{i,1}, \ldots, v_{i,\alpha}$ in I', such that the domain of $v_{i,j}$ is the same as the domain of v_i for every $1 \leq j \leq \alpha$. From now on, we will refer to these variables as "root variables".
2. (permutation variables and permutation metaconstraints) For all $1 \leq i < j \leq n$, we create $\beta * (\beta - 1)$ variables $c_{i,j,1,1}, \ldots, c_{i,j,1,\beta}, c_{i,j,2,1}, \ldots, c_{i,j,2,\beta}, c_{i,j,3,1}, \ldots, c_{i,j,\beta-1,\beta}$ in I' such that the domain of each $c_{i,j,l,col}$ is $\{0, 1, 2, \ldots, 3\beta - 1\}$ for every $1 \leq l \leq \beta - 1$ and every $1 \leq col \leq \beta$. From now on, we will refer to these variables as "permutation variables", and more specifically as the permutation variables "of the permutation metaconstraint between i and j". Informally, the permutation metaconstraint between i and j is a representation in I' of the constraint between the variables v_i and v_j in I. For each $1 \leq l \leq \beta - 1$, we will also refer to the β variables $c_{i,j,l,1}, \ldots, c_{i,j,l,\beta}$ as the "l^{th} line of the permutation metaconstraint between i and j".
Each permutation variable carries two pieces of information: a trilean value to carry to the next line of the permutation metaconstraint, and to which variable of the next line to carry it. The former will be equal to the value assigned to the permutation variable modulo 3. The latter will be determined by dividing the value assigned to the permutation variable by 3, rounding down and adding 1. For example, if the value assigned to $c_{i,j,l,6}$ is 8, then the trilean value 2 will be carried over to $c_{i,j,l+1,3}$ and we therefore know that the value assigned to $c_{i,j,l+1,3}$ will be congruent to 2 modulo 3. One of the permutations will be different from the others: the permutation in the middle of the permutation metaconstraint, from the line $\frac{\beta-1}{2}$ to the line $\frac{\beta-1}{2} + 1$. This permutation will instead represent the transition from the representation of the variable v_i from the original instance I, to the representation of the variable v_j from the original instance I. This particular permutation will be described in the sixth bullet point.
3. (guaranteeing bijections) For all $1 \leq i < j \leq n$, for each $1 \leq l \leq \beta - 1$, for all $1 \leq col < col' \leq \beta$, we introduce the binary constraint $\lfloor \frac{c_{i,j,l,col}}{3} \rfloor \neq \lfloor \frac{c_{i,j,l,col'}}{3} \rfloor$. This is to ensure that in any permutation metaconstraint, the carrying of trilean values from line l to line $l + 1$ follows a bijection.
4. (first line in a permutation metaconstraint) For all $1 \leq i < j \leq n$, for each $1 \leq col \leq \beta$, we set the value 0 in the domain of the root variable $v_{i,col}$ to be incompatible with all values in the domain of the permutation variable $c_{i,j,1,col}$ that are not congruent to 0 modulo 3, we set the value 1 in the domain of $v_{i,col}$ to be incompatible with all values in the domain of $c_{i,j,1,col}$ that are not congruent to 1 modulo 3, and we set the value 2 in the domain of $v_{i,col}$ to be incompatible with all values in the domain of $c_{i,j,1,col}$ that are not congruent to 2 modulo 3.

This is to ensure that $c_{i,j,1,col}$ is the permutation variable in the first line of the permutation metaconstraint between i and j that will carry the trilean value assigned to $v_{i,j}$ to the second line of the permutation metaconstraint between i and j. For example, if the value of $c_{i,j,1,col}$ is 13, it means that the trilean value assigned to v_i, col is 1, and that $c_{i,j,2,5}$ is the permutation variable in the second line of the permutation metaconstraint between i and j that will carry the trilean value assigned to $v_{i,col}$ to the third line of the permutation metaconstraint between i and j.

5. (successive lines in a permutation metaconstraint) For all $1 \leq i < j \leq n$, for each $1 \leq l \leq \beta - 1$ such that $l \neq \frac{\beta-1}{2}$ and $l \neq \beta - 1$, for each $1 \leq col \leq \beta$, for each value a in the domain of $c_{i,j,l,col}$, a will be incompatible with all values in the domain of $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ that are not congruent to a modulo 3.
6. (transition from one original variable to another in the middle of a permutation metaconstraint) For all $1 \leq i < j \leq n$, for each $1 \leq col \leq \beta$, for each value a in the domain of $c_{i,j,\frac{\beta-1}{2},col}$, a will be incompatible with every value b in the domain of $c_{i,j,\frac{\beta-1}{2}+1,\lfloor a/3 \rfloor+1}$ such that ($a \mod 3$) in the domain of v_i is incompatible with ($b \mod 3$) in the domain of v_j in I.
7. (last line in a permutation metaconstraint) For all $1 \leq i < j \leq n$, for all $1 \leq col, col' \leq \beta$, we set the value 0 in the domain of the root variable $v_{j,col'}$ to be incompatible with the values $3(col'-1)+1$ and $3(col'-1)+2$ in the domain of the permutation variable $c_{i,j,\beta-1,col}$, we set the value 1 in the domain of $v_{j,col'}$ to be incompatible with the values $3(col'-1)$ and $3(col'-1)+2$ in the domain of $c_{i,j,\beta-1,col}$, and we set the value 2 in the domain of $v_{j,col'}$ to be incompatible with the values $3(col'-1)$ and $3(col'-1)+1$ in the domain of $c_{i,j,\beta-1,col}$.
8. (preventing contradictions around the root variables) The permutation variables on the first (respectively last) line in a permutation metaconstraint determine the value of the root variables on the previous (respectively following) line, so we need to add the following constraints in order to make sure we do not have a compatible assignment on two permutation variables that implies a contradiction on one of the root variables.

   8.1 For each $1 \leq i \leq n$, for all $1 \leq j_1, j_2 \leq n$ such that $j_1 \neq j_2$, $i < j_1$ and $i < j_2$, for each $1 \leq col \leq \beta$, each value a in the domain of $c_{i,j_1,1,col}$ is incompatible with the values in the domain of $c_{i,j_2,1,col}$ that are not congruent to a modulo 3.

   8.2 For each $1 \leq i \leq n$, for all $1 \leq j_1, j_2 \leq n$ such that $j_1 \neq j_2$, $j_1 < i$ and $j_2 < i$, for all $1 \leq col, col_1, col_2 \leq \beta$, the value $3(col-1)$ in the domain of $c_{j_1,i,\beta-1,col_1}$ is incompatible with the values $3(col-1)+1$ and $3(col-1)+2$ in the domain of $c_{j_2,i,\beta-1,col_2}$, and the value $3(col-1)+1$ in the domain of $c_{j_1,i,\beta-1,col_1}$ is incompatible with the value $3(col-1)+2$ in the domain of $c_{j_2,i,\beta-1,col_2}$.

   8.3 For all $1 \leq j_1 < i < j_2 \leq n$, for all $1 \leq col, col_1 \leq \beta$, the value $3(col-1)$ in the domain of $c_{j_1,i,\beta-1,col_1}$ is incompatible with the values in the domain of $c_{i,j_2,1,col}$ that are not congruent to 0 modulo 3, the value $3(col-1)+1$ in the domain of $c_{j_1,i,\beta-1,col_1}$ is incompatible with the values in the domain of $c_{i,j_2,1,col}$ that are not congruent to 1 modulo 3, and the value $3(col-1)+2$ in the domain of $c_{j_1,i,\beta-1,col_1}$ is incompatible with the values in the domain of $c_{i,j_2,1,col}$ that are not congruent to 2 modulo 3.
9. (the rest) All pairs of values in I' that have not had their compatibility specified yet are set to be allowed.

To illustrate part of the reduction, suppose that we have two variables v_i and v_j in I such that the binary constraint between v_i and v_j is $v_j = (v_i - 1) \mod 3$. Suppose too that $\alpha = 4$, and therefore $\beta = 2\alpha + 1 = 9$. By construction, α cannot be equal to 4, but here we are just choosing a value for α that is small enough to allow an example of reasonable size. Figure 3 represents the permutation metaconstraint between i and j in I'. It is an example of a possible assignment to the 90 variables composing the permutation metaconstraint. The values given to the permutation variables are represented by a couple (a, b), with a being the value to carry to the next line, and b being the variable of the next line to carry to. The dashed arrows represent the transition from v_i to v_j. Note that in this gadget, any compatible assignment on $\alpha = 4$ variables can be extended to a partial solution on all 90 variables.

Let n' be the number of variables in I'. There are $n\beta$ root variables in I'. There is one permutation metaconstraint in I' for each of the $n(n-1)/2$ constraints in I and there are $\beta(\beta-1)$ permutation variables in each permutation metaconstraint of I'. Therefore we have:

$$\begin{aligned}
n' &= n\beta + n(n-1)/2 \times \beta(\beta - 1) \\
&= n(2\alpha+1) + n(n-1)/2 \times 2\alpha(2\alpha+1) \quad \text{(because } \beta = 2\alpha+1\text{)} \\
&= 2n\alpha + n + 2n(n-1)\alpha^2 + n(n-1)\alpha \\
&= 2n(n-1)\alpha^2 + n(n+1)\alpha + n \\
&= 2n^2\alpha^2 - 2n\alpha^2 + n^2\alpha + n\alpha + n \\
&= 2n^8 - 2n^7 + n^5 + n^4 + n \quad \text{(because } \alpha = n^3\text{)} \\
&\leq 2n^8 + n^5 + n^4 + n \\
&\leq 2n^8 + n^8 + n^8 + n^8 \\
&= 5n^8 \\
&\leq n^9 \quad \text{(because } n \geq 12\text{)}
\end{aligned}$$

So the construction of the new instance I' from the original instance I can be done in polynomial time. Furthermore, we have $\alpha = n^3 \geq \sqrt[3]{n'}$. So from Lemma 2, if I' is α^- Minimal then it is $\gamma(n')^-$ Minimal. Therefore, we only have to prove that:

1. if we have a solution for I, then we have a solution for I'.
2. if we have a solution for I', then we have a solution for I.
3. I' is α^- Minimal.

1. if we have a solution for I, then we have a solution for I':
   Let S be a solution for I. For each $1 \leq i \leq n$, let s_i be the trilean value in the domain of v_i belonging to S. Let S' be the following n'-tuple in I':
   (a) (root variables) For each $1 \leq i \leq n$, for each $1 \leq col \leq \beta$, the value s_i in the domain of the root variable $v_{i,col}$ belongs to S'.
   (b) (bottom half of the permutation metaconstraints) For all $1 \leq i < j \leq n$, for each $1 \leq l \leq \alpha$, for each $1 \leq col \leq \beta$, the value $3(col - 1) + s_i$ in the domain of the permutation variable $c_{i,j,l,col}$ belongs to S'.
   (c) (top half of the permutation metaconstraints) For all $1 \leq i < j \leq n$, for each $\alpha + 1 \leq l \leq 2\alpha$, for each $1 \leq col \leq \beta$, the value $3(col - 1) + s_j$ in the domain of the permutation variable $c_{i,j,l,col}$ belongs to S'.

   We are now going to prove that S' is a solution for I'. In order to do so, we have to show that S' satisfies all the constraints in I', which are explicited through bullet points 3 to 8 in the definition of I'.

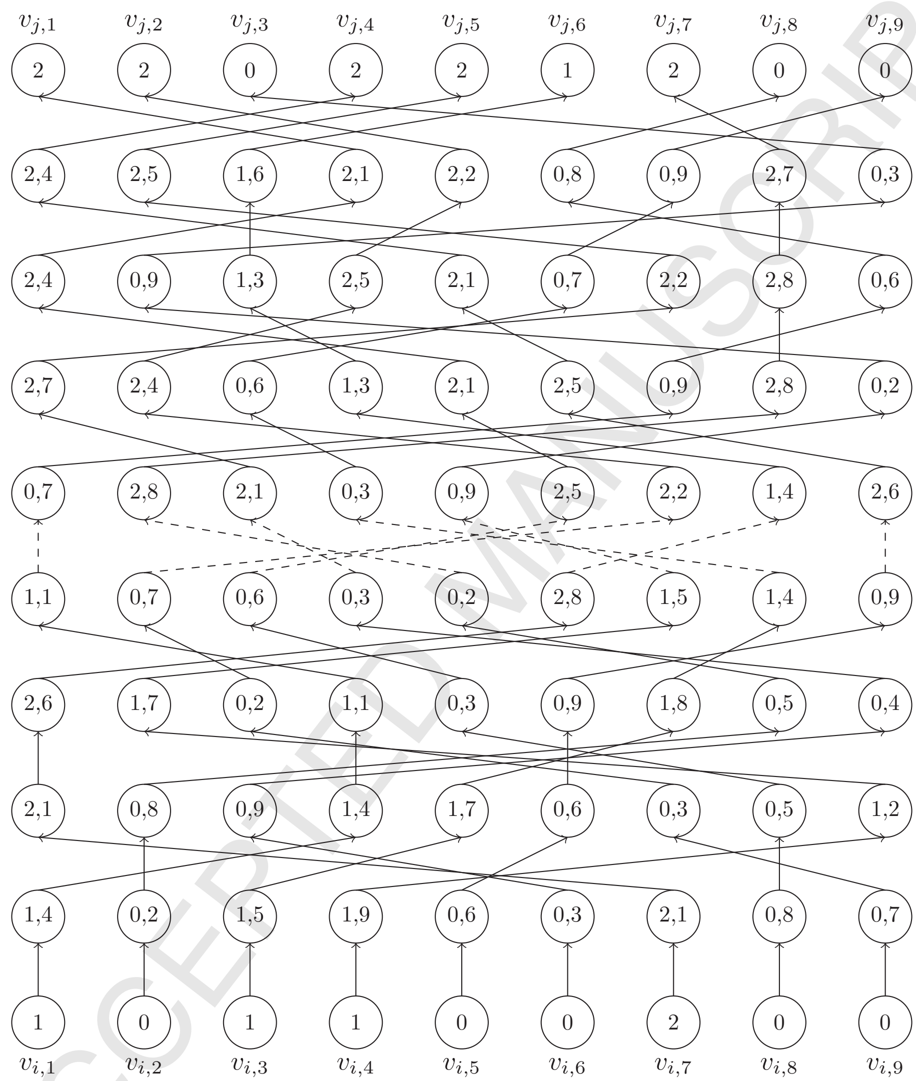

Figure 3: The permutation metaconstraint between i and j.

3. (guaranteeing bijections) Let i and j such that $1 \leq i < j \leq n$, let l such that $1 \leq l \leq \beta - 1$ and let col and col' such that $1 \leq col < col' \leq \beta$. We know from (b) and (c) that the value of S' in the domain of $c_{i,j,l,col}$ is $3(col-1)+a$ and that the value of S' in the domain of $c_{i,j,l,col'}$ is $3(col'-1)+b$, with $0 \leq a, b \leq 2$. So $\lfloor \frac{3(col-1)+a}{3} \rfloor = col - 1 \neq col' - 1 = \lfloor \frac{3(col'-1)+b}{3} \rfloor$, and the constraints given in the third bullet point in the definition of I' are satisfied.
4. (first line in a permutation metaconstraint) Let i and j such that $1 \leq i < j \leq n$ and let col such that $1 \leq col \leq \beta$. We know from (a) and (b) that the value of S' in the domain of $v_{i,col}$ is s_i and that the value of S' in the domain of $c_{i,j,1,col}$ is $3(col-1)+s_i$. Therefore the value of S' in the domain of $v_{i,col}$ is congruent to the value of S' in the domain of $c_{i,j,1,col}$ modulo 3. Therefore the constraints given in the fourth bullet point in the definition of I' are satisfied.
5. (successive lines in a permutation metaconstraint) Let i and j such that $1 \leq i < j \leq n$, let l such that $1 \leq l < \frac{\beta-1}{2}$ and let col such that $1 \leq col \leq \beta$. Let a be the value of S' in the domain of $c_{i,j,l,col}$. From (b), we know that $a = 3(col-1)+s_i$, and that the value of S' in the domain of $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ is also $3(col-1)+s_i$. Therefore the value of S' in the domain of $c_{i,j,l,col}$ is congruent to the value of S' in the domain of $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ modulo 3.

   Similarly, let i and j such that $1 \leq i < j \leq n$, let l such that $\frac{\beta-1}{2} < l < \beta - 1$ and let col such that $1 \leq col \leq \beta$. Let b be the value of S' in the domain of $c_{i,j,l,col}$. From (b), we know that $b = 3(col-1)+s_i$, and that the value of S' in the domain of $c_{i,j,l+1,\lfloor b/3 \rfloor+1}$ is also $3(col-1)+s_i$. Therefore the value of S' in the domain of $c_{i,j,l,col}$ is congruent to the value of S' in the domain of $c_{i,j,l+1,\lfloor b/3 \rfloor+1}$ modulo 3. So the constraints given in the fifth bullet point in the definition of I' are satisfied.
6. (transition from one original variable to another in the middle of a permutation metaconstraint) Let i and j such that $1 \leq i < j \leq n$ and let col such that $1 \leq col \leq \beta$. Let a be the value of S' in the domain of $c_{i,j,\alpha,col}$ and let b be the value of S' in the domain of $c_{i,j,\alpha,\lfloor a/3 \rfloor+1}$. From (b) we know that $(a mod 3) = s_i$ and from (c) we know that $(b mod 3) = s_j$. Since s_i in the domain of v_i and s_j in the domain of v_j are compatible in I, the constraints given in the sixth bullet point in the definition of I' are satisfied.
7. (last line in a permutation metaconstraint) Let i and j such that $1 \leq i < j \leq n$ and let col such that $1 \leq col \leq \beta$. Let a be the value of S' in the domain of $c_{i,j,\beta-1,col}$ and let b be the value of S' in the domain of $v_{j,col}$. From (c) we know that $a = 3(col-1)+s_j$ and from (a) we know that $b = s_j$, so the constraints given in the seventh bullet point in the definition of I' are satisfied.
8. (preventing contradictions around the root variables)

   8.1 Let i, j_1 and j_2 such that $1 \leq i < j_1, j_2 \leq n$ and $j_1 \neq j_2$. Let col such that $1 \leq col \leq \beta$. Let a_1 be the value of S' in the domain of $c_{i,j_1,1,col}$ and let a_2 be the value of S' in the domain of $c_{i,j_2,1,col}$. From (b) we know that $a_1 = a_2 = 3(col-1)+s_i$. So a_1 is congruent to a_2 modulo 3. So the constraints given in bullet point 8.1 in the definition of I' are satisfied.

   8.2 Let i, j_1 and j_2 such that $1 \leq j_1, j_2 < i \leq n$ and $j_1 \neq j_2$. Let col_1 and col_2 such that $1 \leq col_1, col_2 \leq \beta$. Let a_1 be the value of S' in the domain of $c_{j_1,i,\beta-1,col_1}$ and let a_2 be the value of S' in the domain of $c_{j_2,i,\beta-1,col_2}$.

From (c) we know that $a_1 = 3(col_1 - 1) + s_i$ and $a_2 = 3(col_2 - 1) + s_i$. So a_1 is congruent to a_2 modulo 3. So the constraints given in bullet point 8.2 in the definition of I' are satisfied.

8.3 Let i, j_1 and j_2 such that $1 \leq j_1 < i < j_2 \leq n$ and let col and col_1 such that $1 \leq col, col_1 \leq n$. Let a_1 be the value of S' in the domain of $c_{j_1,i,\beta-1,col_1}$ and let a_2 be the value of S' in the domain of $c_{i,j_2,1,col}$. From (c) we know that $a_1 = 3(col_1 - 1) + s_i$ and from b we know that $a_2 = 3(col - 1) + s_i$. So a_1 is congruent to a_2 modulo 3. So the constraints given in bullet point 8.3 in the definition of I' are satisfied.

We have shown that S' satisfies all the constraints in I'. Therefore S' is a solution for I'. So if we have a solution for I, then we have a solution for I'.

2. if we have a solution for I', then we have a solution for I:
Let S' be a solution for I'. For each $1 \leq i \leq n$, for each $1 \leq col \leq \beta$ let $s'_{i,col}$ be the trilean value in the domain of $v_{i,col}$ belonging to S'. For each $1 \leq i \leq n$, let s'_i be the trilean value occuring the most within $\{s'_{i,1}, \ldots, s'_{i,\beta}\}$. Suppose also that for each $1 \leq i \leq n$, s'_i occurs at least $\alpha + 1$ times within $\{s'_{i,1}, \ldots, s'_{i,\beta}\}$.
This last assumption is not trivial, and in order to make it true in the general case, we need to add a gadget, that we call a "majority metaconstraint" to each set of root variables $\{v_{i,1}, \ldots, v_{i,\beta}\}$. Majority metaconstraints are built in a similar way as permutation metaconstraints, so in order to not clutter the proof, we present their construction in Appendix B.
Let S be the n-tuple in I, such that for each $1 \leq i \leq n$, the trilean value s'_i in the domain of v_i belongs to S. We are going to prove that S is a solution for I. In order to do so, we need to show that for all $1 \leq i < j \leq n$, s'_i in the domain of v_i and s'_j in the domain of v_j are compatible in S.
Let i and j such that $1 \leq i < j \leq n$. We have assumed that there are at least $\alpha + 1$ variables within $\{v_{i,1}, \ldots, v_{i,\beta}\}$ that have been assigned the value s'_i by S'. Therefore, from bullet point 4 in the definition of I', we know that there are at least $\alpha+1$ variables on the first line of the permutation metaconstraint between i and j that have been assigned a value congruent to s'_i by S'. From bullet points 3 and 5 in the definition of I', we know that from line 1 to line α in the permutation metaconstraint between i and j, the number of variables in each line that have been assigned a value congruent to s'_i modulo 3 by S' is constant. So by induction we know that there are at least $\alpha + 1$ variables within the $(\frac{\beta-1}{2})^{\text{th}}$ line of the permutation metaconstraint between i and j that have been assigned a value congruent to s'_i modulo 3 by S'. (i)
Similarly, we have assumed that there are at least $\alpha+1$ variables within $\{v_{j,1}, \ldots, v_{j,\beta}\}$ that have been assigned the value s'_j by S'. Therefore, from bullet points 3 and 7 in the definition of I', we know that there are at least $\alpha + 1$ variables on the $(\beta - 1)^{\text{th}}$ of the permutation metaconstraint between i and j that have been assigned a value congruent to s'_j by S'. From bullet points 3 and 5 in the definition of I', we know that from line $\alpha + 1$ to line $\beta - 1$ in the permutation metaconstraint between i and j, the number of variables in each line that have been assigned a value congruent to s'_j modulo 3 by S' is constant. So by induction we know that there are at least $\alpha + 1$ variables within the $(\frac{\beta-1}{2} + 1)^{\text{th}}$ line of the permutation metaconstraint between i and j that have been assigned a value congruent to s'_j modulo 3 by S'. (ii)
From (i), (ii) and bullet point 3 in the definition of I', we know that there is at least

one $1 \leq col \leq \beta$ such that, with a being the value in the domain of $c_{i,j,\alpha,col}$ belonging to S' and b being the value in the domain of $c_{i,j,\alpha+1,\lfloor a/3 \rfloor+1}$ belonging to S', we have both a congruent to s'_i modulo 3 and b congruent to s'_j modulo 3. So from bullet point 6 in the definition of I', we know that s'_i in the domain of v_i and s'_j in the domain of v_j are compatible.
We have shown that S satisfies all the constraints in I. Therefore S is a solution for I. So if we have a solution for I', then we have a solution for I.

3. I' is α^- Minimal:
We need to prove that any partial solution on α or less variables of I' can be extended to a full solution for I'. This part of the proof is extremely long and technical. The complete draft was close to 40 pages. To make it more readable, we only give here the informal idea. Appendix A contains the more detailed version.
The main trick is to notice how all constraints in I' are either between two variables from a same line, or between two variables in two consecutive lines in some metaconstraint. We thus say that two variables are *connected* if they are on the same line or on two consecutive lines in the same metaconstraint. We then partition any partial solution A on α or less variables into k smaller partial solutions $A_1, A_2, \ldots, A_k$ such that:

   - if v and v' are in the same part A_i, then there are some variables $v_1, v_2, \ldots, v_p$ in A_i such that v is connected with v_1, each v_j is connected with v_{j+1} for $1 \leq j < p$ and v_p is connected with v'.
   - if two variables from A are connected, then they are in the same part A_i.

   We show (Lemma 8 in Appendix A) that no two variables from a same part A_i are on the same line. We then extend each part A_i into a partial solution B_i that encompasses exactly one variable on each line. To prove that this is always possible we show that at most two variables of I are represented in each part A_i (Lemma 10 in Appendix A) and use the fact that any pair of two compatible assignments for I can be extended to a full solution for I (because I is minimal).
Each line of I' contains β variables, so to create a full solution for I' it only remains to extend in a similar manner $\beta - k$ empty sets into $\beta - k$ partial solutions that each encompasses exactly one variable on every line. Lemma 14 in Appendix A proves that these partial solutions are compatible with each other and together form the desired full solution.

□

We now generalize the result to any $\epsilon > 2$.

Theorem 5. *Let $\epsilon > 2$ be a real number. Let γ be a function from $\mathbb{N}$ to $\mathbb{R}$ such that $\forall N \in \mathbb{N}, \gamma(N) \leq \sqrt[\epsilon]{N}$. Then the γ^- binary Minimal Constraint Satisfaction Problem is NP-hard.*

Proof. If $\epsilon \geq 3$, then from Proposition 1 and Lemma 2, we have the result. If not, then there is some integer $p > 1$ such that $\epsilon \geq 2 + 1/p$. The proof is then almost the same as the proof of Proposition 1, the only difference being that α is now equal to n^{3p}. The relation between α and β remains the same, with β being equal to $2\alpha + 1$.

Let n' be the number of variables in I'. Let n'_{root} be the number of root variables in I', let n'_{perm} be the number of permutation variables in I', let n'_{orig} be the number of origin variables in I' and let n'_{majo} be the number of majority variables in I'. We have $n' = n'_{root} + n'_{perm} + n'_{orig} + n'_{majo}$. There are $n\beta$ root variables in I', there are $n(n-1)/2$ permutation metaconstraints in I', there are $\beta(\beta-1)$ permutation variables in each permutation metaconstraint of I', there are n majority metaconstraints in I' and there are β origin variables and $(\alpha+1)\beta$ majority variables in each majority metaconstraint of I'. Therefore we have:

$$\begin{aligned}
n' &= n'_{root} + n'_{perm} + n'_{orig} + n'_{majo} \\
&= n\beta + n(n-1)/2 \times \beta(\beta-1) + n\beta + n(\alpha+1)\beta \\
&= n(2\alpha+1) + n(n-1)/2 \times (2\alpha+1)2\alpha + n(2\alpha+1) + n(\alpha+1)(2\alpha+1) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(because } \beta = 2\alpha+1) \\
&= 2n\alpha + n + n(n-1)(2\alpha^2+\alpha) + 2n\alpha + n + n(2\alpha^2+3\alpha+1) \\
&= 2n\alpha + n + n(n-1)(2\alpha^2+\alpha) + 2n\alpha + n + n(2\alpha^2+\alpha) + n(2\alpha+1) \\
&= 2n\alpha + n + n^2(2\alpha^2+\alpha) + 2n\alpha + n + 2n\alpha + n \\
&= 2n^2\alpha^2 + n^2\alpha + 6n\alpha + 3n \\
&= 2n^{2+6p} + n^{2+3p} + 6n^{1+3p} + 3n \quad \text{(because } \alpha = n^{3p}) \\
&\leq 2n^{6p+2} + n^{6p+2} + 6n^{6p+2} + 3n^{6p+2} \\
&= 12n^{6p+2} \\
&\leq n^{6p+3} \quad \text{(because } n \geq 12) \\
&= n^{3p(2+1/p)} \\
&= \alpha^{2+1/p}
\end{aligned}$$

So $\alpha \geq \sqrt[\epsilon]{n'} \geq \gamma(n')$. So if I' is α^- Minimal, then from Lemma 2 I' is α'^- Minimal for each α' such that $2 \leq \alpha' \leq \gamma(n')$. By choosing a high enough value for p, we have the result for all $\epsilon > 2$.

□

5. Conclusion

While Gottlob has proved that the Minimal CSP is NP-hard [6], nothing has been done to determine the limits of its inherent hardness. We presented in this paper three different kinds of expansion for the problem. The first one concerns the size of the domain and the arity constraints. The complexity classification that we established was published [5]. The second one explores the application of minimality to some classes of global constraints. We showed that the hardness of minimality was conserved in this case. The third one, the main result of this paper, generalizes the completability property from compatibilities to a larger set of partial solutions. Again, we proved that no algorithm can exploit in a significant way the additional provided information, unless P=NP.

The work we have done can be extended in a number of different ways. In particular, one could go further down the road of α Minimality, and find out at exactly which point do the classes of instances defined by the α functions transition from NP-hardness to tractability. Because of the close relationship between minimality and consistency, any new discovery about the Minimal CSP might be of considerable help in the study of consistency-based algorithms. Alternatively, one could try and apply the core concept of minimality to problems outside constraint satisfaction, and see whether the properties of minimality are conserved beyond this environment.

Acknowledgements

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289.

[1] Samson Abramsky, Georg Gottlob, and Phokion G. Kolaitis. Robust constraint satisfaction and local hidden variables in quantum mechanics. In Francesca Rossi, editor, *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, pages 440–446. IJCAI/AAAI, 2013.

[2] Jérôme Amilhastre, Hélène Fargier, and Pierre Marquis. Consistency restoration and explanations in dynamic csps application to configuration. *Artif. Intell.*, 135(1-2):199–234, 2002.

[3] Kenneth Appel, Wolfgang Haken, and John Koch. Every planar map is four colorable. II. Reducibility. *Illinois Journal of Mathematics*, 21(3):491–567, 1977.

[4] Christian Bessiere. Constraint propagation. In *Handbook of Constraint Programming*, Foundations of Artificial Intelligence, pages 29–83. Elsevier, 2006.

[5] Guillaume Escamocher and Barry O'Sullivan. On the minimal constraint satisfaction problem: Complexity and generation. In *Combinatorial Optimization and Applications - 9th International Conference, COCOA 2015, Houston, TX, USA, December 18-20, 2015, Proceedings*, pages 731–745, 2015.

[6] Georg Gottlob. On minimal constraint networks. *Artif. Intell.*, 191-192:42–60, 2012.

[7] Lucy Ham. Gap theorems for robust satisfiability: Boolean CSPs and beyond. *Theor. Comput. Sci.*, 676:69–91, 2017.

[8] Ulrich Junker. Configuration. In *Handbook of Constraint Programming*, Foundations of Artificial Intelligence, pages 837–873. Elsevier, 2006.

[9] Ugo Montanari. Networks of constraints: Fundamental properties and applications to picture processing. *Inf. Sci.*, 7:95–132, 1974.

[10] Jean-Charles Régin. A filtering algorithm for constraints of difference in csps. In Barbara Hayes-Roth and Richard E. Korf, editors, *Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 - August 4, 1994, Volume 1.*, pages 362–367. AAAI Press / The MIT Press, 1994.

[11] Francesca Rossi, Peter van Beek, and Toby Walsh. *Handbook of Constraint Programming (Foundations of Artificial Intelligence)*. Elsevier Science Inc., New York, NY, USA, 2006.

[12] Thomas J. Schaefer. The complexity of satisfiability problems. In Richard J. Lipton, Walter A. Burkhard, Walter J. Savitch, Emily P. Friedman, and Alfred V. Aho, editors, *Proceedings of the 10th Annual ACM Symposium on Theory of Computing, May 1-3, 1978, San Diego, California, USA*, pages 216–226. ACM, 1978.

[13] Willem-Jan van Hoeve and Irit Katriel. Global constraints. In *Handbook of Constraint Programming*, Foundations of Artificial Intelligence, pages 169–208. Elsevier, 2006.

Appendix A. Proof of Proposition 1: α^- Minimality of I'

To prove that I' is α^- minimal, we need to show that any compatible tuple of at most α values of I' can be extended to a full solution for I'. The proof is very technical. It requires a large number of specific concepts that we will spend the next few pages defining. The construction of the solution itself is described on Page 30, followed by the necessary lemmas.

Definition 9. We say that two values a and b in I' are ***strongly connected*** if at least one of the following is true:

1. There are some i, j and col such that a is in the domain of the root variable $v_{i,col}$, b is in the domain of the permutation variable $c_{i,j,1,col}$ and b is congruent to a modulo 3.
2. There are some i, j, l, col and col', with $1 \leq l < \beta - 1$ and $l \neq \alpha$, such that a is in the domain of the permutation variable $c_{i,j,l,col}$, b is in the domain of the permutation variable $c_{i,j,l+1,col'}$, $\lfloor a/3 \rfloor + 1 = col'$ and b is congruent to a modulo 3.
3. There are some i, j, col and col' such that a is in the domain of the permutation variable $c_{i,j,\alpha,col}$, b is in the domain of the permutation variable $c_{i,j,\alpha+1,col'}$, $\lfloor a/3 \rfloor + 1 = col'$ and b modulo 3 in the domain of v_j is compatible in I with a modulo 3 in the domain of v_i.
4. There are some i, j, col and col', such that a is in the domain of the permutation variable $c_{i,j,\beta-1,col}$, b is in the domain of the root variable $v_{j,col'}$, $\lfloor a/3 \rfloor + 1 = col'$ and b is congruent to a modulo 3.
5. (ea) There are some i, j_1, j_2 and col, with $j_1 \neq j_2$, such that a is in the domain of the permutation variable $c_{i,j_1,1,col}$, b is in the domain of the permutation variable $c_{i,j_2,1,col}$ and b is congruent to a modulo 3.
   (eb) There are some i, j_1, j_2, col, col_1 and col_2, with $j_1 \neq j_2$, such that a is in the domain of the permutation variable $c_{j_1,i,\beta-1,col_1}$, b is in the domain of the permutation variable $c_{j_2,i,\beta-1,col_2}$, $\lfloor a/3 \rfloor + 1 = \lfloor b/3 \rfloor + 1 = col$ and b is congruent to a modulo 3.
   (ec) There are some i, j_1, j_2, col and col_1 such that a is in the domain of the permutation variable $c_{j_1,i,\beta-1,col_1}$, b is in the domain of the permutation variable $c_{i,j_2,1,col}$, $\lfloor a/3 \rfloor + 1 = col$ and b is congruent to a modulo 3.

Informally, two values a and b are strongly connected if a determines the trilean value carried by b.

Definition 10. We say that two values a and b in I are ***weakly connected***, or simply ***connected***, if at least one of the following is true:

1. $a = b$.
2. a and b are strongly connected.
3. There is a value c in A such that a and c are connected, and b and c are connected.

Definition 11. We say that a set A containing values of I' is a ***connected set*** if all values in A are connected with each other.

Definition 12. Let v be a variable in I. We say that a variable v' in I' is ***associated with*** v if at least one of the following is true:

1. There are some i and col such that v' is the root variable $v_{i,col}$ and v is the variable v_i.
2. There are some i, j, l and col, with $1 \leq l \leq \alpha$, such that v' is the permutation variable $v_{i,j,l,col}$ and v is the variable v_i.
3. There are some i, j, l and col, with $\alpha + 1 \leq l \leq \beta - 1$, such that v' is the permutation variable $v_{i,j,l,col}$ and v is the variable v_i.

We also say that a value a in the domain of a variable v' in I' is associated with a variable v in I if v' is associated with v, and that a set A containing values of I' is associated with a variable v in I if at least one value $a \in A$ is associated with v.

Definition 13. We say that a set V containing β variables of I' is a ***line of variables*** if at least one of the following is true:

1. There is some i such that V is composed of the β root variables $v_{i,1}, \ldots, v_{i,\beta}$.
2. There are some i, j and l such that V is the l^{th} line of the permutation metaconstraint between i and j.

For each $1 \leq i \leq n$, for each $1 \leq j \leq n$ such that $j > i$ and for each $1 \leq l \leq \beta - 1$, we adopt the following conventions:

- L_i is the line of variables composed of the β root variables $v_{i,1}, \ldots, v_{i,\beta}$.
- $L_{i,j,l}$ is the l^{th} line in the permutation metaconstraint between i and j.

Definition 14. We say that two lines of variables L and L' are ***consecutive*** if at least one of the following is true:

1. There are some i and j such that $L = L_i$ and $L' = L_{i,j,1}$.
2. There are some i, j and l, with $1 \leq l \leq \beta - 2$, such that $L = L_{i,j,l}$ and $L' = L_{i,j,l+1}$.
3. There are some i and j such that $L = Li, j, \beta - 1$ and $L' = L_j$.
4. (da) There are some i, j_1 and j_2, with $j_1 \neq j_2$, such that $L = L_{i,j_1,1}$ and $L' = L_{i,j_2,1}$.
   (db) There are some i, j_1 and j_2, with $j_1 \neq j_2$, such that $L = L_{j_1,i,\beta-1}$ and $L' = L_{j_2,i,\beta-1}$.
   (dc) There are some i, j_1 and j_2 such that $L = L_{j_1,i,\beta-1}$ and $L' = L_{i,j_2,1}$.

Note that any non trivial constraint of I' is either between two variables of a same line of variables (bullet point 3 in the definition of I') or between two variables from two consecutive lines of variables (bullet points 4 to 9 in the definition of I'). Note also that if two values a and b of I' are strongly connected, then they are from two consecutive lines of variables.
There are n lines of variables in I' that are composed of root variables, there are $\beta - 1$ lines of permutation variables in each permutation metaconstraint of I' and there are $n(n-1)/2$ permutation metaconstraints in I'. Therefore there are $n + n(n-1)/2 \times (\beta - 1) = n(1 + (n-1)/2 \times 2\alpha) = n(n\alpha - \alpha + 1)$ lines of variables in I'.

Definition 15. We say that a set $V^\star$ of $n(n\alpha - \alpha + 1)$ variables of I' is a ***pseudo-instance of*** I' if $V^\star$ contains exactly one variable from each of the $n(n\alpha - \alpha + 1)$ lines of variables in I'.

Definition 16. Let $V^\star$ be a pseudo-instance of I'. We say that a set $A^\star$ of $n(n\alpha - \alpha + 1)$ values of I' is a ***pseudo-assignment over*** $V^\star$ if all of the following conditions are satisfied:

- For each $v \in V^\star$, exactly one value in the domain of v belongs to $A^\star$.
- For each $v \in V^\star$, for each $v' \in V^\star$ such that v and v' are on consecutive lines of variables, the value of $A^\star$ in the domain of v and the value of $A^\star$ in the domain of v' are strongly connected.

Definition 17. Let $V^\star$ be a pseudo-instance of I'. We say that a set of values $A^\star$ is a ***pseudo-solution of*** $V^\star$ if $A^\star$ is a compatible pseudo-assignment over $V^\star$.

Definition 18. Let a be a value of I', and let v be a variable of I'. We say that v is a ***destination variable of*** a if at least one of the following is true:

1. There are some i, j, and col such that a is in the domain of the root variable $v_{i,col}$ and v is the permutation variable $c_{i,j,1,col}$.
2. There are some i, j, j' and col with $j \neq j'$ such that a is in the domain of the permutation variable $c_{i,j,1,col}$ and v is the permutation variable $c_{i,j',1,col}$.
3. There are some i, j, l and col with $1 \leq l < \beta - 1$ such that a is in the domain of the permutation variable $c_{i,j,l,col}$ and v is the permutation variable $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$.
4. There are some i, j and col such that a is in the domain of the permutation variable $c_{i,j,\beta-1,col}$ and v is the root variable $v_{j,\lfloor a/3 \rfloor+1}$.
5. There are some i_1, i_2, i_3 and col such that a is in the domain of the permutation variable $c_{i_1,i_2,\beta-1,col}$ and v is the permutation variable $c_{i_2,i_3,1,\lfloor a/3 \rfloor+1}$.
6. There are some i, j and col such that a is in the domain of the permutation variable $c_{i,j,1,col}$ and v is the root variable $v_{i,col}$.

Informally, a variable v is a destination variable of a value a if a determines the trilean value carried by v.

Definition 19. Let A be a set of values of I', and let v be a variable of I'. We say that v is a ***destination variable of*** A if there is some value $a \in A$ such that v is a destination variable of A.

Definition 20. Let $A = \{a_1, a_2, \ldots, a_p\}$ be a set of p values of I'. Let $V = \{v_1, v_2, \ldots, v_p\}$ be a set of p variables of I'. We say that V is the ***supporting set of*** A if for each i such that $1 \leq i \leq p$, a_i is in the domain of v_i.

Definition 21. Let v be a variable of I'. We say that the integer col is the ***column of*** v if one the following conditions is fulfilled:

- There is some i such that v is the root variable $v_{i,col}$.
- There are some i, j and l such that v is the permutation variable $c_{i,j,l,col}$.

Definition 22. Let a be a value of I'. We say that the integer col is the ***column of*** a if there is some variable v in I' such that a is in the domain of v and col is the column of v.

Definition 23. Let A be a compatible set of values of I', and let p be such that $1 \le p \le \beta$. We say that the p sets $A_1, A_2, \ldots, A_p$ form a ***connected partition of*** A if all of the following conditions are satisfied:

1. Each value in A appears in exactly one set A_i.
2. If a value a appears in a set A_i, then $a \in A$.
3. For each i with $1 \le i \le p$, A_i is a connected set.
4. For all i and j with $1 \le i < j \le p$, at least one of the following is true:
   - $A_i = \emptyset$.
   - $A_j = \emptyset$.
   - $A_i \cup A_j$ is not a connected set.

Definition 24. Let $A = \{a_1, a_2, \ldots, a_p\}$ be a set of p values from I'. We say that A is ***projectable*** if all of the following conditions are satisfied:

1. For all $1 \le i < j \le p$, a_i and a_j are not from the same line of variables.
2. For all $1 \le i < j \le p$, if a_i and a_j are associated with the same variable v in I, then a_i is congruent to a_j modulo 3.

Definition 25. Let A be a projectable set. Let $A' = \{a'_1, a'_2, \ldots, a'_q\}$ be a set of q values from I. We say that A' is the ***projection of*** A ***over*** I if all of the following conditions are satisfied:

1. For all $1 \le i < j \le q$, a'_i and a'_j are not from the same domain.
2. For each $1 \le i \le q$, there is a value a in A and a variable v in I such that a is associated with v and a'_i is in the domain of v.
3. For each value a in A, there is a value a' in A' and a variable v in I such that a' is in the domain of v, a is associated with v and a is congruent to a' modulo 3.

Definition 26. Let $S = \{s_1, s_2, \ldots, s_n\}$ be a set of n values of I such that s_i is in the domain of v_i for each $1 \le i \le n$. Let $V^\star$ be a pseudo-instance of I'. For each $1 \le i \le n$, let $v_i^\star$ be the variable of $V^\star$ in L_i. For all $1 \le i < j \le n$, for each $1 \le l \le \beta - 1$, let $v^\star_{i,j,l}$ be the variable of $V^\star$ in $L_{i,j,l}$. Let $A^\star$ be a set of values of I' such that each value in $A^\star$ is in the domain of exactly one variable from $V^\star$, and each variable in $V^\star$ has exactly one value from $A^\star$ in its domain. We say that $A^\star$ is the ***projection of*** S ***over*** $V^\star$ if all of the following conditions are satisfied:

1. For each i such that $1 \le i \le n$, s_i is the value of $A^\star$ in the domain of $v_i^\star$.
2. For all i, j and l such that $1 \le i < j \le n$ and $1 \le l \le \alpha$, let col be the column of $v^\star_{i,j,l+1}$. Then $3(col-1) + s_i$ is the value of $A^\star$ in the domain of $v^\star_{i,j,l}$.
3. For all i, j and l such that $1 \le i < j \le n$ and $\alpha + 1 \le l \le \beta - 2$, let col be the column of $v^\star_{i,j,l+1}$. Then $3(col-1) + s_j$ is the value of $A^\star$ in the domain of $v^\star_{i,j,l}$.
4. For all i and j such that $1 \le i < j \le n$, let col be the column of $v_j^\star$. Then $3(col-1) + s_j$ is the value of $A^\star$ in the domain of $v^\star_{i,j,\beta-1}$.

Now that we have the definitions that we need, we explain how to extend a compatible tuple of I' containing at most α values to a solution for I'. All the lemmas needed to prove the correctness of the construction will be stated after. This way, the reader will have already seen the big picture of the proof, and will understand more clearly the purpose of each lemma.

Let A be a compatible α-tuple of I' such that there are at most α values in A. Let $A_1, A_2, \ldots, A_\beta$ be a connected partition of A. From Lemma 3, we know that such a connected partition always exists. From Definition 23(c) and Lemma 5, we know that no two values $a \neq a'$ in a same set A_i are from the same line of variables. Furthermore, from Definition 23(a) and (b), no value a from I' appears in two sets A_i and A_j with $i \neq j$.
Let $V_1, V_2, \ldots, V_\beta$ be the β sets of variables such that for each i, V_i is the supporting set of A_i. Since no two values $a \neq a'$ in a same set A_i are from the same line of variables, no two variables $v \neq v'$ in a same set V_i are from the same line of variables. Furthermore, from Definition 23(a) and (b), no variable v appears in two sets V_i and V_j with $i \neq j$.
Let $V_1^+, V_2^+, \ldots, V_\beta^+$ be the β sets of variables such that for each $1 \leq i \leq \beta$, V_i^+ is the union of V_i and of all the variables in I' that are destination variables of A_i. From Lemma 8, we know that no two variables $v \neq v'$ in a same set V_i^+ are from the same line of variables. From Definition 23 and Lemma 9, we know that no variable v appears in two sets V_i^+ and V_j^+ with $i \neq j$.
Then for each $1 \leq i \leq \beta$ and for each line of variables L of I' that does not contain any variable of V_i^+, we pick a variable v in L such that $\forall j, v \notin V_j^+$, and add v to V_i^+. For each $1 \leq i \leq \beta$, we call $V_i^\star$ the set obtained at the end, when every line of variables of I' contains exactly one variable of $V_i^\star$. From Definition 15, for each i such that $1 \leq i \leq \beta$, $V_i^\star$ is a pseudo-instance of I'. Furthermore, since we made sure that no variable of I' was added to two sets $V_i^\star$ and $V_j^\star$ with $i \neq j$, the β sets $V_1^\star, V_2^\star, \ldots, V_\beta^\star$ form a partition of the set of variables of I'.
For each i such that $1 \leq i \leq \beta$, we know from Lemma 10.1 that A_i is projectable. For each i such that $1 \leq i \leq \beta$, let A'_i be the projection of A over I. For each i such that $1 \leq i \leq \beta$, we know from Lemma 10.2 that A'_i is contains of at most two values. Since I is a Minimal CSP instance, A'_i can be extended to a solution for I for each i such that $1 \leq i \leq \beta$. For each i such that $1 \leq i \leq \beta$, let S_i be a solution for I containing A'_i.
For each i such that $1 \leq i \leq \beta$, let $A_i^\star$ be the projection of S_i over $V_i^\star$. Let $S' = A_1^\star \cup A_2^\star \cup \cdots \cup A_n^\star$. Let a and b be two values of S. If there is some i such that both a and b are in $A_i^\star$, then from Lemma 11 a and b are compatible. Otherwise, let i and j be such that $a \in A_i^\star$ and $b \in A_j^\star$. From Lemma 12, we know that both $A_i^\star$ and $A_j^\star$ are pseudo-assignments. Furthermore, we have already shown that no variable of I' appears in both $V_i^\star$ and $V_j^\star$. So from Lemma 14, a and b are compatible. So S' is a solution for I'.
From Lemma 15, S' contains A. So A can be extended to a solution for I'. So any compatible tuple of I' containing at most α values can be extended to a solution for I'. So I' is α^- Minimal.

Lemma 3. *Let A be a compatible set of at most α values of I', and let p be such that $p \geq |A|$. Then there exists a connected partition $\{A_1, A_2, \ldots, A_q\}$ of A.*

Proof. Let $A = \{a_1, a_2, \ldots, a_q\}$ be a compatible set of $q \leq \alpha$ values of I'. For each i such that $1 \leq i \leq q$, let A_i be the set of values in A that are connected to a_i. If two or more sets are equal, we set all but one of them to be equal to the empty set. If $q < p$, we add $p - q$ empty sets. We need to verify that the final p sets satisfy the conditions from Definition 23:

1. Each value $a_i \in A$ appears in at least the initial set A_i. If A_i is later set to the empty set, it means that there is another set $A_j = A_i$ containing a_i that remains. So each

value in A appears in at least one set A_i. Suppose now that some value $a \in A$ appears in two sets $A_i = A_j$. Without loss of generality, assume that there is a value $b \in A$ such that $b \in A_i$ and $b \notin A_j$. Since both a and b are in A_i, a and b are connected. Since a is in A_j, a is connected with a_j. Since connectivity is a transitive property, b is also connected with a_j. So b should be in A_j and we have a contradiction. So each value in A appears in at most one set A_i.

2. All the values in the initial sets are values from A. Then no new value is added, the only other operations are removing values and adding empty sets. So for each i such that $1 \leq i \leq p$, for each $a \in A_i$, a belongs to A.
3. Let i be such that $1 \leq i \leq p$. Either A_i is one of the initial sets, and therefore there is a value $a_j \in A_i$ such that all values of A_i are connected to a_j, or A_i is an empty set. In either case, A_i is a connected set.
4. Suppose that we have two non-empty sets $A_i \neq A_j$ such that $A_i \cup A_j$ is a connected set. Without loss of generality, assume that there is a value $b \in A$ such that $b \in A_i$ and $b \notin A_j$. Since $b \in A_i \cup A_j$ and $A_i \cup A_j$ is a connected set, b is connected with all values in $A_i \cup A_j$. In particular, b is connected with A_j. So b should be in A_j and we have a contradiction. So the union of two non-empty sets A_i and A_j with $i \neq j$ is not a connected set.

□

We are now going to map the lines of variables of I', using the "consecutive" relation from Definition 14. In order to do this we introduce G, a graph defined by the following:

- The lines of variables of I' are the vertices of G.
- Two lines of variables of I' are connected in G if they are consecutive.

From Definition 14(a), (c) and (d), for each i such that $1 \leq i \leq n$, the n lines of variables $L_{1,i,\beta-1}, L_{2,i,\beta-1}, \ldots, L_{i-1,i,\beta-1}, L_i, L_{i,i+1,1}, \ldots, L_{i,n,1}$ form a clique C_i in G. There are n such cliques, one for each variable in I. From Definition 14(b), the rest of the edges in G is composed of the paths $P_{i,j}$ of the $\beta - 3$ lines of variables $L_{i,j,2}, L_{i,j,3}, \ldots, L_{i,j,\beta-2}$ between each pair of cliques (C_i, C_j) with $i < j$.

The graph G, as well as its cliques C_i and paths $P_{i,j}$, will be used in the proofs of several lemmas below.

Lemma 4. *Let $P = \{e_1, e_2, \ldots, e_p\}$ be a path of p distinct lines of variables in G. Let i be such that both e_1 and e_p are in C_i. Then for each q, e_q is in C_i.*

Proof. Suppose that at least one line of variables in P is not in C_i. Let e_q be the first line of variables to leave C_i, that is $e_{q'}$ is in C_i for each $1 \leq q' < q$. So e_q is part of a path $P_{j,i}$ (respectively $P_{i,j}$). So $e_q = L_{j,i,\beta-2}$ (respectively $e_q = L_{i,j,2}$) and $e_{q-1} = L_{j,i,\beta-1}$ (respectively $e_{q-1} = L_{i,j,1}$). Since there are at most α line of variables in P and there are $\beta - 3 > \alpha$ lines of variables in $P_{j,i}$ (respectively $P_{i,j}$), P cannot reach the clique C_j. So if $e_{q+1} = L_{j,i,\beta-3}$ (respectively $e_{q+1} = L_{i,j,3}$), then P would have to pass a second time through $L_{j,i,\beta-2}$ (respectively $L_{i,j,2}$) to go back to C_i, which is not possible because we assumed that all lines of variables in P are distinct. From Definition 14, the only line of variables other than $L_{i,j,3}$ (respectively other than $L_{j,i,\beta-3}$) that is consecutive with $L_{i,j,2}$ (respectively with $L_{j,i,\beta-2}$) is $L_{i,j,1}$ (respectively $L_{j,i,\beta-1}$). But if $e_{q+1} = L_{j,i,\beta-1}$

(respectively if $e_{q+1} = L_{i,j,1}$), then $e_{q-1} = e_{q+1}$, which contradicts the conditions of the lemma. Therefore, there is no $e_q \in P$ such that $e_q \notin C_i$ and we have the desired result. □

Lemma 5. *Let A be a compatible connected set with at most α values. Let a and b be two values from A. If $a \neq b$, then a and b are not from the same line of variables.*

Proof. Let $P = \{e_1, e_2, \ldots, e_p\}$ be a path of p lines of variables in G, such that $2 \leq p \leq \alpha$, $e_1, e_2, \ldots, e_{p-1}$ are all different and $e_1 = e_p$. There are two possibilities for e_1:

- e_1 is part of a clique C_i. Suppose that there is a line of variables of P that is not part of the clique C_i. Let e_q be the first line of variables of P that is not in C_i, by that we mean that $e_{q'}$ is in C_i for each $q' < q$. So e_q is part of a path $P_{j,i}$ (respectively $P_{i,j}$). So $e_q = L_{j,i,\beta-2}$ (respectively $e_q = L_{i,j,2}$) and $e_{q-1} = L_{j,i,\beta-1}$ (respectively $e_{q-1} = L_{i,j,1}$). Since there are at most α line of variables in P, there are $\beta - 3 > \alpha$ lines of variables in $P_{j,i}$ (respectively $P_{i,j}$) and all lines of variables in P are different with the exception of $e_1 = e_p$, P cannot reach the clique C_j. So if $e_{q+1} = L_{j,i,\beta-3}$ (respectively $e_{q+1} = L_{i,j,3}$), then P would have to pass a second time through $L_{j,i,\beta-2}$ (respectively $L_{i,j,2}$) to go back to C_i, which is not possible because we assumed that the only repeated line of variables in P is in C_i. So $e_{q+1} = L_{j,i,\beta-1}$ (respectively $e_{q+1} = L_{i,j,1}$) and P only contains the three lines of variables $\{e_1 = L_{j,i,\beta-1}, e_2 = L_{j,i,\beta-2}, e_3 = L_{j,i,\beta-1}\}$ (respectively $\{e_1 = L_{i,j,1}, e_2 = L_{i,j,2}, e_3 = L_{i,j,1}\}$). So if e_1 is in a clique C_i, then the only three possibilities are:
  - There is some j such that $P = \{L_{j,i,\beta-1}, L_{j,i,\beta-2}, L_{j,i,\beta-1}\}$.
  - There is some j such that $P = \{L_{i,j,1}, L_{i,j,2}, L_{i,j,1}\}$.
  - All of the lines of variables in P are in C_i.
- e_1 is part of a path $P_{i,j}$. So there is l with $2 \leq l \leq \beta - 2$ such that $e_1 = e_p = L_{i,j,l}$. So $e_2 = L_{i,j,l-1}$ (respectively $e_2 = Li, j, l+1$). From there, e_3 can be equal to $L_{i,j,l-2}$, to L_i if $l = 2$, or to $L_{i,j,l}$ (respectively $L_{i,j,l+2}$, L_j if $l = \beta - 2$ and $L_{i,j,l}$). In the first two cases, P cannot pass through $L_{i,j,l-1}$ (respectively $L_{i,j,l+1}$) again because we assumed that $L_{i,j,l}$ is the only repeated line of variables in P. So P has to go to the clique C_i (respectively C_j) through the lines of variables $L_{i,j,l-1}, L_{i,j,l-2}, \ldots, L_{i,j,3}, L_{i,j,2}$ (respectively $L_{i,j,l-1}, L_{i,j,l-2}, \ldots, L_{i,j,3}, L_{i,j,2}$), then it has to go to the clique C_j (respectively C_i) without passing through $P_{i,j}$, then it has to go through the lines of variables $L_{i,j,\beta-2}, L_{i,j,\beta-3}, \ldots, L_{i,j,l+2}, L_{i,j,l+1}$ (respectively $L_{i,j,2}, L_{i,j,3}, \ldots, L_{i,j,l-2}$, $L_{i,j,l-1}$) in order to go back to $L_{i,j,l}$. Since there are at most α lines of variables in P and the length of each path is $\beta - 3$, this is not possible. So if e_1 is in a path $P_{i,j}$, then the only two possibilities are:
  - There is some l such that $P = \{L_{i,j,l}, L_{i,j,l-1}, L_{i,j,l}\}$.
  - There is some l such that $P = \{L_{i,j,l}, L_{i,j,l+1}, L_{i,j,l}\}$.

Let a_0 and b be two values of A, such that $a_0 \neq b$ and a_0 and b are from the same line of variables. From the definition of strongly connected, we know that two values can only be strongly connected if they are from two consecutive (and therefore distinct) lines of variables. So a and b are not strongly connected. So a and b are weakly connected. So there are p values $a_1, a_2, \ldots, a_p$ in A, with $1 \leq p \leq \alpha - 2$, such that for all $0 \leq i < j \leq \alpha - 2$,

$a_i \neq a_j$, $\forall 0 \leq i < p$, a_i is strongly connected to a_{i+1}, and a_p is strongly connected to b. If two values a_i and a_j with $0 \leq i < j \leq p$ are from the same line of variables, we set a_0 to be equal to a_i, b to be equal to a_j, p to be equal to $j - i - 1$, and each a_q with $1 \leq q \leq p$ to be equal to (the former) a_{i+q}. We repeat this operation until all values within $\{a_0, a_1, \ldots, a_p\}$ are from different lines of variables. We can only do this operation a finite number of times, because the value of p strictly decreases every time we do the operation. Furthermore, p is always greater or equal to 1, because two strongly connected values cannot be on the same line of variables. Therefore, we have a path $P = \{a_0, a_1, \ldots, a_p, b\}$ of $p+2$ different values such that a_i and a_{i+1} are from consecutive lines of variables for each $0 \leq i < p$, a_p and b are from consecutive lines of variables, a_0 and b are from the same line of variables and all values within $\{a_0, a_1, \ldots, a_p\}$ are from different lines of variables. As we have previously shown, there are only five possibilities in regard to the supporting set of P.

- $P = \{a_0, a_1, b\}$ and there are some i, j, col_1, col_2 and col_3 such that a_0 is in the domain of the permutation variable $c_{j,i,\beta-1,col_1}$, a_1 is in the domain of the permutation variable $c_{j,i,\beta-2,col_2}$ and b is in the domain of the permutation variable $c_{j,i,\beta-1,col_3}$. The only bullet point in the definition of strongly connected (Definition 9) that applies to the last two lines of constraints in a permutation metaconstraint is (b). From Definition 9(b), $col_1 = col_3 = \lfloor a_1/3 \rfloor + 1$. So a_0 and b are from the same domain, and therefore cannot be part of the compatible set A.

- $P = \{a_0, a_1, b\}$ and there are some i, j, col_1, col_2 and col_3 such that a_0 is in the domain of the permutation variable $c_{j,i,1,col_1}$, a_1 is in the domain of the permutation variable $c_{j,i,2,col_2}$ and b is in the domain of the permutation variable $c_{j,i,1,col_3}$. The only bullet point in the definition of strongly connected (Definition 9) that applies to the last two lines of constraints in a permutation metaconstraint is (b). From Definition 9(b), $\lfloor a_0/3 \rfloor + 1 = \lfloor b/3 \rfloor + 1 = col_2$. So from the third bullet point in the definition of I', a_0 and b are either in the same domain or incompatible. In either case, this contradicts the conditions of the lemma.

- There is some i such that all the values in P are from the lines of variables of the clique C_i. The only bullet points in the definition of strongly connected (Definition 9) that apply to two lines of variables in a same clique C_i are (a), (d) and (e). For each a in P, either there is some j such that a is from the line of variables $L_{j,i,\beta-1}$ or there are some j and col such that a is in the domain of the root variable $v_{i,col}$ or a is in the domain of the permutation variable $c_{i,j,1,col}$. Let $f(a)$ be equal to $\lfloor a/3 \rfloor + 1$ in the former case and let $f(a)$ be equal to col in the latter case. Let a and a' be two strongly connected values in P. From Definition 9, in particular the bullet points (a), (d) and (e), $f(a) = f(a')$. Therefore, there is some col such that $f(a) = col$ for each a in P. Therefore, $f(a_0) = f(b)$. So if a_0 and b are both from the line of variables L_i, or if there is some j such that a_0 and b are both from the line of variables $L_{i,j,1}$, then a_0 and b are from the same domain, and therefore cannot be part of the compatible set A. So there is some j such that a_0 and b are both from the line of variables $L_{j,i,\beta-1}$, and $\lfloor a_0/3 \rfloor + 1 = \lfloor b/3 \rfloor + 1 = col$. So from the third bullet point in the definition of I', a_0 and b are either in the same domain or incompatible. In either case, this contradicts the conditions of the lemma.

- $P = \{a_0, a_1, b\}$ and there are some i, j, $2 \leq l \leq \beta - 2$, col_1, col_2 and col_3 such that

a_0 is in the domain of the permutation variable c_{i,j,l,col_1}, a_1 is in the domain of the permutation variable $c_{i,j,l-1,col_2}$ and b is in the domain of the permutation variable c_{i,j,l,col_3}. The only bullet points in the definition of strongly connected (Definition 9) that apply to two successive lines of permutation constraints in a permutation meta-constraint are (b) and (c). From Definition 9(b) and (c), $col_1 = col_3 = \lfloor a_1/3 \rfloor + 1$. So a_0 and b are from the same domain, and therefore cannot be part of the compatible set A.

- $P = \{a_0, a_1, b\}$ and there are some i, j, $2 \leq l \leq \beta - 2$, col_1, col_2 and col_3 such that a_0 is in the domain of the permutation variable c_{i,j,l,col_1}, a_1 is in the domain of the permutation variable $c_{i,j,l+1,col_2}$ and b is in the domain of the permutation variable c_{i,j,l,col_3}. The only bullet points in the definition of strongly connected (Definition 9) that apply to two successive lines of permutation constraints in a permutation meta-constraint are (b) and (c). From Definition 9(b) and (c), $\lfloor a_0/3 \rfloor + 1 = \lfloor b/3 \rfloor + 1 = col_2$. So from the third bullet point in the definition of I', a_0 and b are either in the same domain or incompatible. In either case, this contradicts the conditions of the lemma.

□

Lemma 6. *Let A be a connected set with at most α values. Let a and a' be two values of A. Let i be such that both a and a' are from the clique C_i. For each value $b \in A$ from C_i, let $f(b)$ be equal to $\lfloor b/3 \rfloor + 1$ if there is some j such that b is from the line of variables $L_{j,i,\beta-1}$ and $f(b)$ be equal to the column of b otherwise. Then $f(a) = f(a')$.*

Proof. From Definition 10(b) and (c), we know that there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ of $p + 2$ values in A, with $0 \leq p \leq \alpha - 2$, such that a is strongly connected to a_1, a_q is strongly connected to a_{q+1} for each $1 \leq q \leq p$ and a_p is strongly connected to a'. From Lemma 5, we know that all the values in P are from different lines of variables. So from Lemma 4, we know that all the values in P are from C_i. Let b and b' be two strongly connected values in P. From Definition 9, in particular the bullet points (a), (d) and (e), $f(b) = f(b')$. Therefore, there is some col such that $f(b) = col$ for each $b \in P$, which completes the proof. □

Lemma 7. *Let A be a connected set containing at most α values. Let a and a' be two values of A, such that a and a' are from consecutive lines. Then a and a' are strongly connected.*

Proof. Let a and a' be two values of A, such that a is from the line of variables L, a' is from the line of variables L' and L and L' are consecutive. So the relation between L and L' follows one of the conditions from Definition 14.

1. The relation between L and L' follows Definition 14(a). So there are some i and j such that $L = L_i$ and $L' = L_{i,j,1}$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a and a', with $0 \leq p \leq \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \leq i < p$ and a_p is strongly connected to a'. If $p = 0$ we have the result, so we can assume without loss of generality that $p > 0$. From Lemma 4, we know that all the values in P are from C_i. Therefore from Definition 9, in particular the bullet points (a), (d) and (e), all the values in P are congruent to each other modulo 3. In particular, a is congruent to a' modulo 3. Furthermore, from Lemma 6, there is some col such that a is in the

domain of the root variable $v_{i,col}$ and a' is in the domain of the root variable $c_{i,j,1,col}$. So from Definition 9(a), a and a' are strongly connected.

2. The relation between L and L' follows Definition 14(b). So there are some i, j and l with $1 \le l \le \beta - 2$ such that $L = L_{i,j,l}$ and $L' = L_{i,j,l+1}$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a and a', with $0 \le p \le \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \le i < p$ and a_p is strongly connected to a'. Suppose that $p > 0$. Since $1 \le l \le \beta - 2$, we know that either $l \ge 2$ or $l \le \beta - 3$. Suppose that $l \ge 2$ (respectively suppose that $l \le \beta - 3$). Then a_1 is either from $L_{i,j,l-1}$ or from $L_{i,j,l+1}$ (respectively a_p is either from $L_{i,j,l+2}$ or from $L_{i,j,l}$). Suppose first that a_1 (respectively a_p) is from $L_{i,j,l-1}$ (respectively $L_{i,j,l+2}$). Since a is from $L_{i,j,l}$ (respectively a' is from $L_{i,j,l+1}$, from Lemma 4 no value from $\{a_2, a_3, \ldots, a_p, a'\}$ (respectively $\{a, a_1, a_2, \ldots, a_{p-2}, a_{p-1}\}$) can be from $L_{i,j,l}$ (respectively $L_{i,j,l+1}$). So in order to go to $L_{i,j,l+1}$ (respectively $L_{i,j,l}$) from a_1 (respectively a_p), P has to go to C_i (respectively C_j), then to C_j (respectively C_i) without passing by the path $P_{i,j}$, then through the lines of variables $L_{i,j,\beta-2}, L_{i,j,\beta-3}, \ldots, L_{i,j,l+1}$ (respectively $L_{i,j,2}, L_{i,j,3}, \ldots, L_{i,j,l}$). Since there are at most α values in P, this is not possible. So a_1 (respectively a_p) is not from $L_{i,j,l-1}$ (respectively $L_{i,j,l+2}$). So a_1 (respectively a_p) is from $L_{i,j,l+1}$ (respectively $L_{i,j,l}$). So a_1 (respectively a_p) is from the same line of variables as a' (respectively a). From Lemma 5, this is not possible. So $p = 0$ and we have the result.

3. The relation between L and L' follows Definition 14(c). So there are some i and j such that $L = L_{i,j,\beta-1}$ and $L' = L_j$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a and a', with $0 \le p \le \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \le i < p$ and a_p is strongly connected to a'. If $p = 0$ we have the result, so we can assume without loss of generality that $p > 0$. From Lemma 4, we know that all the values in P are from C_i. Therefore from Definition 9, in particular the bullet points (a), (d) and (e), all the values in P are congruent to each other modulo 3. In particular, a is congruent to a' modulo 3. Furthermore, from Lemma 6, there is some *col* such that $\lfloor a/3 \rfloor + 1 = col$ and a' is in the domain of the root variable $v_{j,col}$. So from Definition 9(d), a and a' are strongly connected.

4. (da) The relation between L and L' follows Definition 14(da). So there are some i, j_1 and j_2 with $j_1 \ne j_2$ such that $L = L_{i,j_1,1}$ and $L' = L_{i,j_2,1}$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a and a', with $0 \le p \le \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \le i < p$ and a_p is strongly connected to a'. If $p = 0$ we have the result, so we can assume without loss of generality that $p > 0$. From Lemma 4, we know that all the values in P are from C_i. Therefore from Definition 9, in particular the bullet points (a), (d) and (e), all the values in P are congruent to each other modulo 3. In particular, a is congruent to a' modulo 3. Furthermore, from Lemma 6, there is some *col* such that a is in the domain of the permutation variable $c_{i,j_1,1,col}$ and a' is in the domain of the permutation variable $c_{i,j_2,1,col}$. So from Definition 9(ea), a and a' are strongly connected.

   (db) The relation between L and L' follows Definition 14(db). So there are some i, j_1 and j_2 with $j_1 \ne j_2$ such that $L = L_{j_1,i,\beta-1}$ and $L' = L_{j_2,i,\beta-1}$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a

and a', with $0 \leq p \leq \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \leq i < p$ and a_p is strongly connected to a'. If $p = 0$ we have the result, so we can assume without loss of generality that $p > 0$. From Lemma 4, we know that all the values in P are from C_i. Therefore from Definition 9, in particular the bullet points (a), (d) and (e), all the values in P are congruent to each other modulo 3. In particular, a is congruent to a' modulo 3. Furthermore, from Lemma 6, there is some col such that $\lfloor a/3 \rfloor + 1 = \lfloor a'/3 \rfloor + 1 = col$. So from Definition 9(eb), a and a' are strongly connected.

(dc) The relation between L and L' follows Definition 14(dc). So there are some i, j_1 and j_2 such that $L = L_{j_1,i,\beta-1}$ and $L' = L_{i,j_2,1}$. From Definition 10(b) and (c) there is a path $P = \{a, a_1, a_2, \ldots, a_p, a'\}$ between a and a', with $0 \leq p \leq \alpha - 2$, such that a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \leq i < p$ and a_p is strongly connected to a'. If $p = 0$ we have the result, so we can assume without loss of generality that $p > 0$. From Lemma 4, we know that all the values in P are from C_i. Therefore from Definition 9, in particular the bullet points (a), (d) and (e), all the values in P are congruent to each other modulo 3. In particular, a is congruent to a' modulo 3. Furthermore, from Lemma 6, there is some col such that $\lfloor a/3 \rfloor + 1 = col$ and a' is in the domain of the permutation variable $c_{i,j_2,1,col}$. So from Definition 9(ec), a and a' are strongly connected.

□

Lemma 8. *Let A be a compatible connected set with at most α values. Let V be the union of the supporting set of A with the set of variables of I' that are destination variables of A. Let v and v' be two variables in V, such that $v \neq v'$. Then v and v' are not from the same line of variables.*

Proof. First, note that from Definition 18, if a is a value of I' and L is a line of variables of I', then there is at most one destination variable of a in L.
Let v and v' be two variables of V such that v and v' are from the same line of variables L. If neither v nor v' is a destination variable of A, then from Lemma 5 $v = v'$. Otherwise, we have two possibilities:

1. There is a value $a \in A$ in the domain of v and v' is a destination variable of A. Let a' be a value of A such that v' is a destination variable of a'. There are three possibilities for L:

   - There is some i such that $L = L_i$. So from Definition 18, in particular bullet points (d) and (f), there is some j such that a' is from $L_{j,i,\beta-1}$ (respectively $L_{i,j,1}$). So from Definition 14(c) (respectively Definition 14(a)), a and a' are from consecutive lines. So from Lemma 7, they are strongly connected. So from Definition 9, in particular bullet point (d) (respectively (a)), $\lfloor a'/3 \rfloor + 1$ is equal to the column of a (respectively the column of a' is equal to the column of a). So from Definition 18, in particular bullet point (d) (respectively (f)), a is in the domain of a destination variable of a'. Since there is at most one destination variable of a' in L_i, $v = v'$.

   - There are some i and j such that $L = L_{i,j,1}$. So from Definition 18, in particular bullet points (a), (b) and (e), there is some j' such that a' is from L_i (respectively

$L_{i,j',1}$ and $L_{j',i,\beta-1}$). So from Definition 14(a) (respectively Definition 14(da) and Definition 14(dc)), a and a' are from consecutive lines. So from Lemma 7, they are strongly connected. So from Definition 9, in particular bullet point (a) (respectively (ea) and (ec)), the column of a' is equal to the colum of a (respectively the column of a' is equal to the column of a and $\lfloor a'/3 \rfloor + 1$ is equal to the column of a). So from Definition 18, in particular bullet point (a) (respectively (b) and (e)), a is in the domain of a destination variable of a'. Since there is at most one destination variable of a' in $L_{i,j,1}$, $v = v'$.

- There are some i, j and $l > 1$ such that $L = L_{i,j,l}$. So from Definition 18, in particular bullet point (c), a' is from $L_{i,j,l-1}$. So from Definition 14(b), a and a' are from consecutive lines. So from Lemma 7, they are strongly connected. So from Definition 9, in particular bullet points (b) and (c), $\lfloor a'/3 \rfloor + 1$ is equal to the column of a. So from Definition 18, in particular bullet point (c), a is in the domain of a destination variable of a'. Since there is at most one destination variable of a' in $L_{i,j,l}$, $v = v'$.

2. Both v and v' are destination variables of A. Let a and a' be two values of A such that v is a destination variable of a and v' is a destination variable of a'. Let L_a be the line of variable of a and let $L_{a'}$ be the line of variables of a'. There are three possibilities for L:

   - There is some i such that $L = L_i$. So from Definition 18, in particular bullet points (d) and (f), there are some j_1, j_2, j'_1 and j'_2 such that a is from $L_{j_1,i,\beta-1}$ or $L_{i,j_2,1}$ and a' is from $L_{j'_1,i,\beta-1}$ or $L_{i,j'_2,1}$. Furthermore, from Lemma 5, a and a' are not from the same line of variables. So from Definition 14(d), a and a' are from consecutive lines. So from Lemma 7, a and a' are strongly connected. Let $f(a)$ be equal to $\lfloor a/3 \rfloor + 1$ if a is from $L_{j_1,i,\beta-1}$ and let $f(a)$ be equal to the column of a if a is from $L_{i,j_2,1}$. Similarly, let $f(a')$ be equal to $\lfloor a'/3 \rfloor + 1$ if a' is from $L_{j'_1,i,\beta-1}$ and let $f(a')$ be equal to the column of a' if a' is from $L_{i,j'_2,1}$. From Lemma 6, $f(a) = f(a')$. So from Definition 18(d) and (f), $v_{i,f(a)}$ is a destination variable of both a and a'. Since there is at most one destination variable of a' in L_i, $v = v'$.

   - There are some i and j such that $L = L_{i,j,1}$. So from Definition 18, in particular bullet points (a), (b) and (e), there are some j_1, j_2, j'_1 and j'_2 with $j_2 \neq j$ and $j'_2 \neq j$ such that a is from $L_{j_1,i,\beta-1}$, L_i or $L_{i,j_2,1}$ and a' is from $L_{j'_1,i,\beta-1}$, L_i or $L_{i,j'_2,1}$. Furthermore, from Lemma 5, a and a' are not from the same line of variables. So from Definition 14(a), (c) and (d), a and a' are from consecutive lines. So from Lemma 7, a and a' are strongly connected. Let $f(a)$ be equal to $\lfloor a/3 \rfloor + 1$ if a is from $L_{j_1,i,\beta-1}$ and let $f(a)$ be equal to the column of a if a is from L_i or $L_{i,j_2,1}$. Similarly, let $f(a')$ be equal to $\lfloor a'/3 \rfloor + 1$ if a' is from $L_{j'_1,i,\beta-1}$ and let $f(a')$ be equal to the column of a' if a' is from L_i or $L_{i,j'_2,1}$. From Lemma 6, $f(a) = f(a')$. So from Definition 18(a), (b) and (e), $c_{i,j,1,f(a)}$ is a destination variable of both a and a'. Since there is at most one destination variable of a' in $L_{i,j,1}$, $v = v'$.

   - There are some i, j and l with $l > 1$ such that $L = L_{i,j,l}$. So from Definition 18, in particular bullet point c, both a and a' are from $L_{i,j,l-1}$. So from Lemma 5,$a = a'$. Therefore $v = v'$.

We have shown that if two variables of V are in the same line of variables, then they are equal. This completes the proof. □

Lemma 9. *Let A and A' be two compatible sets of values, with at most α values in each set, such that the union of A and A' is a compatible set, the union of A and A' is not a connected set and $\forall a \in A, \forall a' \in A'$, a and a' are not in the same domain. Let v be a destination variable of A and let v' be a destination variable of A'. Then $v \neq v'$.*

Proof. Suppose that $v = v'$. Then there is a value $a \in A$ and a value $a' \in A'$ such that v is a destination variable of both a and a'. If a and a' are connected, then $A \cup A'$ is a connected set from Definition 10(c), so the conditions of the lemma are not satisfied. There are three possibilities for v:

1. There are some i, j and col such that v is the permutation variable $c_{i,j,1,col}$. Then the relation between v on one hand, and a and a' on the other hand, follows (a), (b) or (e) in Definition 18. Therefore, there are 9 possibilities for the combination of the two relations:
   (a) The relation between a and v follows Definition 18(a) and the relation between a' and v follows Definition 18(a). From Definition 18(a), it means that both a and a' are in the domain of the root variable $v_{i,col}$, which is in contradiction with the conditions of the lemma.
   (b) The relation between a and v follows Definition 18(a) and the relation between a' and v follows Definition 18(b). From Definition 18(a) and (b), it means that a is in the domain of the root variable $v_{i,col}$ and that there is a j' such that a' is in the domain of the permutation variable $c_{i,j',1,col}$. Either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(a), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 4 in the definition of I'. In any case, the conditions of the lemma are not satisfied.
   (c) The relation between a and v follows Definition 18(a) and the relation between a' and v follows Definition 18(e). From Definition 18(a) and (e), it means that a is in the domain of the root variable $v_{i,col}$ and that there are some j' and col' such that a' is in the domain of the permutation variable $c_{j',i,\beta-1,col'}$ and $\lfloor a'/3 \rfloor + 1 = col$. Either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(d), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 7 in the definition of I'. In any case, the conditions of the lemma are not satisfied.
   (d) The relation between a and v follows Definition 18(b) and the relation between a' and v follows Definition 18(a). Same argument as in (ii) by switching the roles of a and a'.
   (e) The relation between a and v follows Definition 18(b) and the relation between a' and v follows Definition 18(b). From Definition 18(b), it means that there are some j_1 and j_2 such that a is in the domain of the permutation variable $c_{i,j_1,1,col}$ and a' is in the domain of the permutation variable $c_{i,j_2,1,col}$. If $j_1 = j_2$, then a and a' are in the same domain and the conditions of the lemma are not satisfied. Otherwise, either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(ea), or a is not congruent to a' modulo 3, in

which case a and a' are incompatible from bullet point 8.1 in the definition of I'. In any case, the conditions of the lemma are not satisfied.

(f) The relation between a and v follows Definition 18(b) and the relation between a' and v follows Definition 18(e). From Definition 18(b) and (e), it means that there are some j_1, j_2 and col' such that a is in the domain of the permutation variable $c_{i,j_1,1,col}$, a' is in the domain of the permutation variable $c_{j_2,i,\beta-1,col'}$ and $\lfloor a'/3 \rfloor + 1 = col$. Either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(ec), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 8.3 in the definition of I'. In any case, the conditions of the lemma are not satisfied.

(g) The relation between a and v follows Definition 18(e) and the relation between a' and v follows Definition 18(a). Same argument as in (iii) by switching the roles of a and a'.

(h) The relation between a and v follows Definition 18(e) and the relation between a' and v follows Definition 18(b). Same argument as in (vi) by switching the roles of a and a'.

(i) The relation between a and v follows Definition 18(e) and the relation between a' and v follows Definition 18(e). From Definition 18(e), it means that there are some j_1, j_2, col_1 and col_2 such that a is in the domain of the permutation variable $c_{j_1,i,\beta-1,col_1}$, a' is in the domain of the permutation variable $c_{j_2,i,\beta-1,col_2}$ and $\lfloor a/3 \rfloor + 1 = \lfloor a'/3 \rfloor + 1 = col$. If $j_1 = j_2$, either a and a' are in the same domain, or they are incompatible from bullet point 3 in the definition of I'. If $j_1 \neq j_2$, either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(eb), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 8.2 in the definition of I'. In any case, the conditions of the lemma are not satisfied.

2. There are some i, j, l and col with $1 \leq l < \beta - 1$ such that v is the permutation variable $c_{i,j,l+1,col}$. Then the relation between v on one hand, and a and a' on the other hand, follows (c) in Definition 18. From Definition 18(c), there are some col_1 and col_2, such that a is in the domain of the permutation variable c_{i,j,l,col_1}, a' is in the domain of the permutation variable c_{i,j,l,col_1} and $\lfloor a/3 \rfloor + 1 = \lfloor a/3 \rfloor + 1 = col$. If a and a' are in the same domain, then the conditions of the lemma are not satisfied. If a and a' are not from the same domain, then from bullet point 3 in the definition of I', a and a' are incompatible, which is in contradiction with the conditions of the lemma.

3. There are some j and col such that v is the root variable $v_{j,col}$. Then the relation between v on one hand, and a and a' on the other hand, follows either (d) or (f) in Definition 18. Therefore, there are 4 possibilities for the combination of the two relations:

(a) The relation between a and v follows Definition 18(d) and the relation between a' and v follows Definition 18(d). From Definition 18(d), there are some i_1, i_2, col_1 and col_2 such that a is in the domain of the permutation variable c_{i_1,j,l,col_1}, a' is in the domain of the permutation variable c_{i_2,j,l,col_2} and $\lfloor a/3 \rfloor + 1 = \lfloor a'/3 \rfloor + 1 = col$. If a and a' are in the same domain, then the conditions of the lemma are not satisfied. If a and a' are not from the same domain, then either $i_1 = i_2$, in which case a and a' are incompatible from bullet point 3 in the definition of I', or $i_1 \neq i_2$. In the latter case, either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(eb), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 8.2 in the definition of I'. In any case, the conditions of the lemma are not satisfied.

(b) The relation between a and v follows Definition 18(d) and the relation between a' and v follows Definition 18(f). From Definition 18(d) and (f), there are some i_1, i_2 and col' such that a is in the domain of the permutation variable $c_{i,i_1,1,col}$, a' is in the domain of the permutation variable $c_{i_2,i,\beta-1,col'}$ and $\lfloor a'/3 \rfloor + 1 = col$. Either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(ec), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 8.3 in the definition of I'. In any case, the conditions of the lemma are not satisfied.

(c) The relation between a and v follows Definition 18(f) and the relation between a' and v follows Definition 18(d). Same argument as in (ii) by switching the roles of a and a'.

(d) The relation between a and v follows Definition 18(f) and the relation between a' and v follows Definition 18(f). From Definition 18(f), there are some i_1, i_2 and col such that a is in the domain of the permutation variable $c_{j,i_1,1,col}$ and a' is in the domain of the permutation variable $c_{j,i_2,1,col}$. If $i_1 = i_2$, then a and a' are in the same domain and the conditions of the lemma are not satisfied. Otherwise, either a is congruent to a' modulo 3, in which case a and a' are connected from Definition 9(ea), or a is not congruent to a' modulo 3, in which case a and a' are incompatible from bullet point 8.1 in the definition of I'. In any case, the conditions of the lemma are not satisfied and the proof of the lemma is complete.

□

Lemma 10. *Let A be a connected set with at most α values. Then:*

1. *A is projectable.*
2. *The projection of A over I contains at most two values.*

Proof. Let $P = \{a_1, a_2, \ldots, a_p\}$ be a path of p values in A such that a_i is strongly connected to a_{i+1} for each $1 \leq i < p$. Suppose that there are some q, i_1 and i_2 such that a_q is associated with v_{i_1}, a_{q+1} is associated with v_{i_2} and $v_{i_1} \neq v_{i_2}$. From Definition 9, in particular bullet point (c), a_q is from $L_{i_1,i_2,\alpha}$ (respectively $L_{i_2,i_1,\alpha+1}$) and a_{q+1} is from $L_{i_1,i_2,\alpha+1}$ (respectively $L_{i_2,i_1,\alpha}$). We are going to prove that a_{q+r} is in $L_{i_1,i_2,\alpha+r}$ (respectively $L_{i_2,i_1,\alpha+1-r}$) for each $0 \leq r \leq p-q$. We have shown that the property is true for $r = 0$ and $r = 1$. Suppose that the property is true for each r such that $0 \leq r \leq r'$, with $1 \leq r' < p-q$. Then $a_{q+r'-1}$ is in $L_{i_1,i_2,\alpha+r'-1}$ (respectively $L_{i_2,i_1,\alpha+1-r'+1}$) and $a_{q+r'}$ is in $L_{i_1,i_2,\alpha+r'}$ (respectively $L_{i_2,i_1,\alpha+1-r'}$). So $a_{q+r'+1}$ is either in $L_{i_1,i_2,\alpha+r'-1}$ (respectively $L_{i_2,i_1,\alpha+1-r'+1}$) or in $L_{i_1,i_2,\alpha+r'+1}$ (respectively $L_{i_2,i_1,\alpha+1-r'-1}$). But from Lemma 5, $a_{q+r'+1}$ cannot be in the same line of variables as $a_{q+r'-1}$. So $a_{q+r'+1}$ is in $L_{i_1,i_2,\alpha+r'+1}$ (respectively $L_{i_2,i_1,\alpha+1-r'-1}$). So the property is true for each r such that $0 \leq r \leq r'+1$, with $1 \leq r' < p-q$. So by induction the property is true for each r such that $0 \leq r \leq p-q$. So from Definition 12, $a_{q'}$ is associated with v_{i_2} for each q' such that $q < q' \leq p$.

1. We need to prove that both conditions in Definition 24 are satisfied.
   (a) From Lemma 5.

(b) Let a and b be two values of A such that a and b are associated with the same variable v of I. From the definition of a connected set, we know that a and b are connected. So from Definition 10(b) and (c), there is a path of values $P = \{a, a_1, a_2, \ldots, a_p, b\}$ from a to b, with $0 \leq p \leq \alpha - 2$, such that each a is strongly connected to a_1, a_i is strongly connected to a_{i+1} for each $1 \leq i \leq p$ and a_p is strongly connected to b. If there is some q such that a_q is associated with $v' \neq v$, then we have shown in the preamble of the proof that $a_q, a_{q+1}, a_{q+2}, \ldots, a_p, b$ are all associated with $v' \neq v$. This is not possible because we assumed that b is associated with v. So all values in P are associated with v. So from Definition 9(a), (b), (d) and (e), all values in P are congruent to each other modulo 3. In particular, a is congruent to b modulo 3.

2. The number of values in the projection of A over I is the number of variables v_i in I such that at least one value in A is associated with v_i. Let a_0, b_0 and c be three different values in A. Let v_i, v_j and v_h be the variables in I such that a_0 is associated with v_i, b_0 is associated with v_j and c is associated with v_h. Suppose that i, j and h are all different. Without loss of generality, assume that $i < j < h$. Since A is a connected set, there exists a path $P = \{a_0, a_1, a_2, \ldots, a_p, b_0, b_1, b_2, \ldots, b_q, c\}$ of $p+q+3$ values of A, with $p \leq \alpha-2$ and $q \leq \alpha-2$, such that a_i is strongly connected to a_{i+1} for each $0 \leq i < p$, a_p is strongly connected to b_0, b_i is strongly connected to b_{i+1} for each $1 \leq i < q$ and b_q is strongly connected to c. As we have shown in the preamble in the proof, there are some p' and q' with $p' \geq 1$ and $q' \geq 1$ such that the values $a_0, a_1, a_2, \ldots, a_{p'-1}$ are associated with v_i, the values $a_{p'}, a_{p'+1}, a_{p'+2}, \ldots, a_p, b_0, b_1, b_2, \ldots, b_{q'-1}$ are associated with v_j and the values $b_{q'}, b_{q'+1}, b_{q'+2}, \ldots, b_q, c$ are associated with v_h. So $a_{p'-1}$ and $a_{p'}$ are from $L_{i,j,\alpha}$ and $L_{i,j,\alpha+1}$ respectively, the two lines of variables in the middle of the path $P_{i,j}$, while $b_{q'-1}$ and $b_{q'}$ are from $L_{j,h,\alpha+1}$ and $L_{j,h,\alpha}$ respectively, the two lines of variables in the middle of the path $P_{j,h}$. So P has to go through $L_{i,j,\alpha}, L_{i,j,\alpha+1}, L_{i,j,\alpha+2}, \ldots, L_{i,j,\beta-2}$, then through at least one line of variables from C_j, then through $L_{j,h,\beta-2}$, $L_{j,h,\beta-3}, \ldots, L_{j,h,\alpha+1}, L_{j,h,\alpha}$, for a total of at least $(\beta-2-\alpha+1)+1+(\beta-2-\alpha+1) = 2\alpha+1-2-\alpha+1+1+2\alpha+1-2-\alpha+1 = 2\alpha$ variables. Since the number of values in P is $p+q+3 \leq (\alpha-2)+(\alpha-2)+3 = 2\alpha-1$, this is not possible. So no three values of A are associated with three different variables in I. So there are at most two values in the projection of A over I.

□

Lemma 11. *Let S be a solution for I. Let $V^\star$ be a pseudo-instance of I' such that for all i and j with $1 \leq i < j \leq n$, the column of the variable of $V^\star$ in L_i is equal to the column of the variable of $V^\star$ in $L_{i,j,1}$. Then the projection of S over $V^\star$ is compatible.*

Proof. For each i such that $1 \leq i \leq n$, let s_i be the value of S in the domain of v_i. Let $A^\star$ be the projection of S over $V^\star$. To show that $A^\star$ is compatible, we just have to show that $A^\star$ satisfies all the constraints in I'. These constraints are given through bullet points 3 to 8 in the definition of I':

3. These constraints only apply to values within a same line of variables. From Definition 15, no line of variables of I' contains two different variables of $V^\star$. Therefore the constraints given in the third bullet point in the definition of I' are always satisfied.

4. Let i and j be such that $1 \leq i < j \leq n$. Let a be the value of $A^\star$ in L_i and let b be the value of $A^\star$ in $L_{i,j,1}$. From the conditions of the Lemma, we know that the column of a is equal to the column of b. From Definition 26(a) and (b) respectively, we know that both a and b are congruent to s_i modulo 3. So from the fourth bullet point in the definition of I', a and b are compatible. Therefore the constraints given in the fourth bullet point in the definition of I' are always satisfied.
5. From Definition 26(b) and (c).
6. Let i and j be such that $1 \leq i < j \leq n$. Let a be the value of $A^\star$ in $L_{i,j,\alpha}$ and let b be the value of $A^\star$ in $L_{i,j,\alpha+1}$. From Definition 26(b), we know that a is congruent to s_i modulo 3, and from Definition 26(c) we know that b is congruent to s_j modulo 3. Since S is a solution for I, s_i in the domain of v_i and s_j in the domain of v_j are compatible. Furthermore, from Definition 26(b), $\lfloor a/3 \rfloor + 1$ is equal to the column of b. So from the sixth bullet point in the definition of I', a and b are compatible. Therefore the constraints given in the sixth bullet point in the definition of I' are always satisfied.
7. From Definition 26(d).
8.
   8.1 Let i, j and j' be such that $1 \leq i < j, j' \leq n$ and $j \neq j'$. Let a be the value of $A^\star$ in $L_{i,j,1}$ and let b be the value of $A^\star$ in $L_{i,j',1}$. From Definition 26(b), we know that both a and b are congruent to s_i modulo 3. Furthermore, we know from the conditions of the lemma that both the column of a and the column of b are equal to the value of $A^\star$ in L_i. So from bullet point 8.1 in the definition of I', a and b are compatible. Therefore the constraints given in bullet point 8.1 in the definition of I' are always satisfied.

   8.2 From Definition 26(d).

   8.3 Let i, j and j' be such that $1 \leq j < i < j' \leq n$. Let a be the value of $A^\star$ in $L_{j,i,\beta-1}$ and let b be the value of $A^\star$ in $L_{i,j',1}$. From Definition 26(d) and (b) respectively, we know that both a and b are congruent to s_i modulo 3. From Definition 26(d), we also know that $\lfloor a/3 \rfloor + 1$ is equal to the column of the value of $A^\star$ in L_i. From the conditions of the lemma we know that the column of b is also equal to the value of $A^\star$ in L_i. So from bullet point 8.3 in the definition of I', a and b are compatible. Therefore the constraints given in bullet point 8.3 in the definition of I' are always satisfied.

□

Lemma 12. *Let S be a solution for I. Let $V^\star$ be a pseudo-instance of I' such that for all i and j with $1 \leq i < j \leq n$, the column of the variable of $V^\star$ in L_i is equal to the column of the variable of $V^\star$ in $L_{i,j,1}$. Then the projection of S over $V^\star$ is a pseudo-assignment.*

Proof. Let $A^\star$ be the projection of S over $V^\star$. Definition 26 explicitely states that there is exactly one value of $A^\star$ in the domain of each variable from $V^\star$, so $A^\star$ satisfies the first bullet point in the definition of a pseudo-assignment (Definition 16). Therefore, we only have to prove that $A^\star$ satisfies the second bullet point in Definition 16, that is that any couple of values in $A^\star$ from consecutive lines are strongly connected.
Let a and b be two values of $A^\star$ such that a and b are from consecutive lines. From Definition 14, we know that there are six possibilities for the relation between a and b:

1. There are some i and j such that a is from L_i and b is from $L_{i,j,1}$. From Definition 26(a), we know that a is equal to s_i. From Definition 26(b), we know that b

is congruent to s_i modulo 3. From the conditions of the current lemma, we know that the columns of a and b are equal. So from Definition 9(a), a and b are strongly connected.

2. There are some i, j and l with $1 \le l \le \beta - 2$ such that a is from $L_{i,j,l}$ and b is from $L_{i,j,l+1}$. Let *col* be the column of b. Depending on the exact value of l, there are three possibilities:
   - $1 \le l < \alpha$: From Definition 26(b), we know that a is equal to $3(col - 1) + s_i$ and that b is congruent to s_i modulo 3. So from Definition 9(b), a and b are strongly connected.
   - $l = \alpha$: From Definition 26(b), we know that a is equal to $3(col - 1) + s_i$. From Definition 26(c), we know that a is congruent to s_j modulo 3. Since S is a solution for I, s_i is the domain of v_i and s_j in the domain of v_j are compatible. So from Definition 9(c), a and b are strongly connected.
   - $\alpha + 1 \le l \le \beta - 2$: From Definition 26(c), we know that a is equal to $3(col-1)+s_j$ and that b is congruent to s_j modulo 3. So from Definition 9(b), a and b are strongly connected.
3. There are some i and j such that a is from $L_{i,j,\beta-1}$ and b is from L_j. Let *col* be the column of b. From Definition 26(d), we know that a is equal to $3(col - 1) + s_j$. From Definition 26(a), we know that b is equal to s_j. So from Definition 9(d), a and b are strongly connected.

(da) There are some i, j_1 and j_2 with $j_1 \ne j_2$ such that a is from $L_{i,j_1,1}$ and b is from $L_{i,j_2,1}$. From Definition 26(b), we know that both a and b are congruent to s_i modulo 3. From the conditions of the current lemma, we know that the columns of a and b are equal. So from Definition 9(ea), a and b are strongly connected.

(db) There are some i, j_1 and j_2 with $j_1 \ne j_2$ such that a is from $L_{j_1,i,\beta-1}$ and b is from $L_{j_2,i,\beta-1}$. Let *col* be the column of the value of $A^\star$ in L_i. From Definition 26(d), we know that both a and b are equal to $3(col - 1) + s_i$. So from Definition 9(eb), a and b are strongly connected.

(dc) There are some i, j_1 and j_2 with $j_1 \ne j_2$ such that a is from $L_{j_1,i,\beta-1}$ and b is from $L_{i,j_2,1}$. Let *col* be the column of the value of $A^\star$ in L_i. From Definition 26(d), we know that a is equal to $2(col - 1) + s_i$. From the conditions of the current lemma, we know that the column of b is *col*. So from Definition 9(ec), a and b are strongly connected.

In all cases a and b are strongly connected, so we have the result. □

Lemma 13. *Let $V^\star$ be a pseudo-instance of I'. Let $A^\star$ be a pseudo-assignment over $V^\star$. Let a be a value of $A^\star$. Let v be a destination variable of a. Then $v \in V^\star$.*

Proof. Let a' be the value of A from the same line of variables as v. Definition 18 gives six different possibilities for the relation between a and v.

- If this relation follows the situation laid out in Definition 18(a), then there are some i, j and *col* such that a is in the domain of the root variable $v_{i,col}$ and v is the permutation variable $c_{i,j,1,col}$. So from Definition 14(a) we know that a and a' are from consecutive lines of variables, namely L_i and $L_{i,j,1}$. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet point in Definition 9 that allows two values from the lines of variables L_i and $L_{i,j,1}$ to be strongly connected is (a). So from Definition 9(a), a' is in the domain of the variable $c_{i,j,1,col}$, which is actually v. So $v \in V^\star$.

- If this relation follows the situation laid out in Definition 18(b), then there are some i, j, j' and *col* with $j \neq j'$ such that a is in the domain of the permutation variable $c_{i,j,1,col}$ and v is the permutation variable $c_{i,j',1,col}$. So from Definition 14(da) we know that a and a' are from consecutive lines of variables, namely $L_{i,j,1}$ and $L_{i,j',1}$. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet point in Definition 9 that allows two values from the lines of variables $L_{i,j,1}$ and $L_{i,j',1}$ to be strongly connected is (ea). So from Definition 9(ea), a' is in the domain of the variable $c_{i,j',1,col}$, which is actually v. So $v \in V^\star$.

- If this relation follows the situation laid out in Definition 18(c), then there are some i, j, l and *col* with $1 \leq l < \beta - 1$ such that a is in the domain of the permutation variable $c_{i,j,l,col}$ and v is the permutation variable $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$. So from Definition 14(b) we know that a and a' are from consecutive lines of variables, namely $L_{i,j,l}$ and $L_{i,j,l+1}$. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet points in Definition 9 that allow two values from the lines of variables $L_{i,j,l}$ and $L_{i,j,l+1}$ to be strongly connected are (b) and (c). So from both Definition 9(b) and Definition 9(c), a' is in the domain of the variable $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$, which is actually v. So $v \in V^\star$.

- If this relation follows the situation laid out in Definition 18(d), then there are some i, j and *col* such that a is in the domain of the permutation variable $c_{i,j,\beta-1,col}$ and v is the root variable $v_{j,\lfloor a/3 \rfloor+1}$. So from Definition 14(c) we know that a and a' are from consecutive lines of variables, namely $L_{i,j,\beta-1}$ and L_j. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet point in Definition 9 that allows two values from the lines of variables $L_{i,j,\beta-1}$ and L_j to be strongly connected is (d). So from Definition 9(d), a' is in the domain of the variable $v_{j,\lfloor a/3 \rfloor+1}$, which is actually v. So $v \in V^\star$.

- If this relation follows the situation laid out in Definition 18(e), then there are some i_1, i_2, i_3 and *col* such that a is in the domain of the permutation variable $c_{i_1,i_2,\beta-1,col}$ and v is the permutation variable $c_{i_2,i_3,1,\lfloor a/3 \rfloor+1}$. So from Definition 14(dc) we know that a and a' are from consecutive lines of variables, namely $L_{i_1,i_2,\beta-1}$ and $L_{i_2,i_3,1}$. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet point in Definition 9 that allows two values from the lines of variables $L_{i_1,i_2,\beta-1}$ and $L_{i_2,i_3,1}$ to be strongly connected is (ec). So from Definition 9(ec), a' is in the domain of the variable $c_{i_2,i_3,1,\lfloor a/3 \rfloor+1}$, which is actually v. So $v \in V^\star$.

- If this relation follows the situation laid out in Definition 18(f), then there are some i, j and *col* such that a is in the domain of the permutation variable $c_{i,j,1,col}$ and v is the root variable $v_{i,col}$. So from Definition 14(a) we know that a and a' are from consecutive lines of variables, namely L_i and $L_{i,j,1}$. So from the second bullet point in the definition of a pseudo-assignment, we know that a and a' are strongly connected. The only bullet point in Definition 9 that allows two values from the lines of variables L_i and $L_{i,j,1}$ to be strongly connected is (a). So from Definition 9(a), a' is in the domain of the variable $v_{i,col}$, which is actually v. So $v \in V^\star$.

□

Lemma 14. *Let $V_1^\star, V_2^\star, \ldots, V_\beta^\star$ be β pseudo-instances of I' such that for each variable v of I', v appears in exactly one of the β pseudo-instances. For each p such that $1 \leq p \leq \beta$, let $A_p^\star$ be a pseudo-assignment over the domains of the pseudo-instance $V_p^\star$. Let p and q be such that $1 \leq p < q \leq \beta$, and let a and b be two values such that $a \in A_p^\star$ and $b \in A_q^\star$. Then a and b are compatible.*

Proof. Suppose that a and b are incompatible. Therefore a and b violate one of the constraints of I'. These constraints are explicited through bullet points 3 to 8 in the definition of I'.

3. There are some i and j with $1 \leq i < j \leq n$, some l with $1 \leq l \leq \beta - 1$, some col and col' with $1 \leq col < col' \leq \beta$, such that a is in the domain of the permutation variable $c_{i,j,l,col}$, b is in the domain of the permutation variable $c_{i,j,l,col'}$ and $\lfloor a/3 \rfloor = \lfloor b/3 \rfloor$. So if $l < \beta - 1$, then from Definition 18(c), the permutation variable $c_{i,j,l+1,\lfloor a/3 \rfloor}$ is a destination variable of both a and b and if $l = \beta - 1$, then from Definition 18(d) the root variable $v_{j,\lfloor a/3 \rfloor}$ is a destination variable of both a and b. In either case, there is a variable v of I' such that v is a destination variable of both a and b. So from Lemma 13, v belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of the third bullet point in the definition of I'.
4. There are some i, j and col with $1 \leq i < j \leq n$ and $1 \leq col \leq \beta$ such that a is in the domain of the root variable $v_{i,col}$ and b is in the domain of the permutation variable $c_{i,j,1,col}$. So $c_{i,j,1,col}$ belongs to $V_q^\star$. But from Definition 18(a), $c_{i,j,1,col}$ is a destination variable of a. So from Lemma 13, $c_{i,j,1,col}$ belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of the fourth bullet point in the definition of I'.

5. and 6. There are some i, j, l and col with $1 \leq i < j \leq n$, $1 \leq l < \beta - 1$ and $1 \leq col \leq \beta$ such that a is in the domain of the permutation variable $c_{i,j,l,col}$ and b is in the domain of the permutation variable $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$. So $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ belongs to $V_q^\star$. But from Definition 18(c), $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ is a destination variable of a. So from Lemma 13, $c_{i,j,l+1,\lfloor a/3 \rfloor+1}$ belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of the fifth or sixth bullet point in the definition of I'.

7. There are some i, j and col with $1 \leq i < j \leq n$ and $1 \leq col \leq \beta$ such that a is in the domain of the permutation variable $c_{i,j,\beta-1,col}$ and b is in the domain of the root variable $v_{j,\lfloor a/3 \rfloor+1}$. So $v_{j,\lfloor a/3 \rfloor+1}$ belongs to $V_q^\star$. But from Definition 18(d), $v_{j,\lfloor a/3 \rfloor+1}$ is a destination variable of a. So from Lemma 13, $v_{j,\lfloor a/3 \rfloor+1}$ belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of the seventh bullet point in the definition of I'.
8. 8.1 There are some i, j_1, j_2 and col with $1 \leq i < j_1, j_2 \leq n$, $j_1 \neq j_2$ and $1 \leq col \leq \beta$ such that a is in the domain of the permutation variable $c_{i,j_1,1,col}$ and b is in the

domain of the permutation variable $c_{i,j_2,1,col}$. So $c_{i,j_2,1,col}$ belongs to $V_q^\star$. But from Definition 18(b), $c_{i,j_2,1,col}$ is a destination variable of a. So from Lemma 13, $c_{i,j_2,1,col}$ belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of bullet point 8.1 in the definition of I'.

8.2 There are some i, j_1, j_2, col, col_1 and col_2 with $1 \leq j_1, j_2 < i \leq n$, $j_1 \neq j_2$ and $1 \leq col, col_1, col_2 \leq \beta$ such that a is in the domain of the permutation variable $c_{j_1,i,\beta-1,col_1}$, b is in the domain of the permutation variable $c_{j_2,i,\beta-1,col_2}$ and $\lfloor a/3 \rfloor = \lfloor b/3 \rfloor = col$. So from Definition 18(d) the root variable $v_{i,col}$ is a destination variable of both a and b. So there is a variable v of I' such that v is a destination variable of both a and b. So from Lemma 13, v belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of bullet point 8.2 in the definition of I'.

8.3 There are some j_1, i, j_2, col and col_1 with $1 \leq j_1 < i < j_2 \leq n$ and $1 \leq col, col_1 \leq \beta$ such that a is in the domain of the permutation variable $c_{j_1,i,\beta-1,col_1}$ and b is in the domain of the permutation variable $c_{i,j_2,1,\lfloor a/3+1 \rfloor}$. So $c_{i,j_2,1,\lfloor a/3+1 \rfloor}$ belongs to $V_q^\star$. But from Definition 18(e), $c_{i,j_2,1,\lfloor a/3+1 \rfloor}$ is a destination variable of a. So from Lemma 13, $c_{i,j_2,1,\lfloor a/3+1 \rfloor}$ belongs to both $V_p^\star$ and $V_q^\star$, which contradicts the conditions of the lemma. So the incompatibility between a and b cannot be because of bullet point 8.3 in the definition of I'.

We have shown that no constraint of I' can be violated by a and b. This proves the lemma. □

Lemma 15. *Let A_0 be a projectable set of values of I'. Let a be a value in A_0. Let S be a solution for S containing the projection of A_0 over I. Let $V_i^\star$ be a pseudo-instance of I' such that $V_i^\star$ contains all destination variables of a. Let $A_i^\star$ be the projection of S over $V_i^\star$. Then a belongs to $A_i^\star$.*

Proof. For each i such that $1 \leq i \leq n$, let s_i be the value of S in the domain of v_i. There are four possibilities for a:

1. There are some i and col such that a is in the domain of the root variable $v_{i,col}$. From Definition 25(c), we know that s_i is equal to a. So from Definition 26(a), a belongs to $A_i^\star$.
2. There are some i, j, l and col with $1 \leq l \leq \alpha$ such that a is in the domain of the permutation variable $c_{i,j,l,col}$. From Definition 25(c), we know that a is congruent to s_i modulo 3. Let col' be equal to $\lfloor a/3 \rfloor + 1$. From Definition 18(c), $c_{i,j,l+1,col'}$ is a destination variable of a. Since $V_i^\star$ contains all destination variables of a, $c_{i,j,l+1,col'} \in V_i^\star$. So from Definition 26(b), a belongs to $A_i^\star$.
3. There are some i, j, l and col with $\alpha + 1 \leq l \leq \beta - 2$ such that a is in the domain of the permutation variable $c_{i,j,l,col}$. From Definition 25(c), we know that a is congruent to s_j modulo 3. Let col' be equal to $\lfloor a/3 \rfloor + 1$. From Definition 18(c), $c_{i,j,l+1,col'}$ is a destination variable of a. Since $V_i^\star$ contains all destination variables of a, $c_{i,j,l+1,col'} \in V_i^\star$. So from Definition 26(c), a belongs to $A_i^\star$.
4. There are some i, j and col such that a is in the domain of the permutation variable $c_{i,j,\beta-1,col}$. From Definition 25(c), we know that a is congruent to s_j modulo 3. Let col' be equal to $\lfloor a/3 \rfloor + 1$. From Definition 18(d), $v_{j,col'}$ is a destination variable of a.

Since $V_i^\star$ contains all destination variables of a, $v_{j,col'} \in V_i^\star$. So from Definition 26(d), a belongs to $A_i^\star$.

In all cases, a belongs to $A_i^\star$ and we have the result. □

Appendix B. Proof of Proposition 1: Majority Metaconstraints

In the proof of Proposition 1, we need to show that we can build a solution S to the original instance I from a solution S' to the constructed instance I'. To do so, for each i between 1 and n, we look at the $2\alpha+1$ values assigned to the $2\alpha+1$ variables $v_{i,1}, \ldots, v_{i,2\alpha+1}$ in I', and pick the value among 0, 1 and 2 that occurs at least $\alpha + 1$ times. Since each variable can be assigned one of three different values, and there are only $2\alpha + 1$ variables, it is not clear why there will always be one value represented at least $\alpha + 1$ times. To make sure this is always the case, we build a "majority metaconstraint" M_i for each $1 \leq i \leq n$ that will ensure absolute majority while keeping α^- minimality.
Here is the description of M_i for a given $1 \leq i \leq n$:

1. (origin variables) A set of β variables $w_{i,1}, \ldots, w_{i,\beta}$. Each one of these β variables has a domain of size 3 containing the three values 0, 1 and 2. From now on, we will refer to these β variables as "origin variables".
2. (enforcing a majority) For all $1 \leq j \neq j' \leq \alpha+1$, we add an equality constraint between the two origin variables $w_{i,j}$ and $w_{i,j'}$. This is the original part of the instance where the property of the lemma will be enforced.
3. (majority variables and majority metaconstraints) A set of $\beta(\alpha+1)$ variables $m_{i,1,1}, \ldots, m_{i,1,\beta}, m_{i,2,1}, \ldots, m_{i,2,\beta}, m_{i,3,1}, \ldots, m_{i,\alpha+1,\beta}$. From now on, we will refer to these variables as "majority variables". For each $1 \leq l \leq \alpha + 1$, we will also refer to the β variables $m_{i,l,1}, \ldots, m_{i,l,\beta}$ as the "l^{th} line of the majority metaconstraint M_i". Each of the $\beta(\alpha+1)$ majority variables has a domain containing $3 \times \beta \times 6$ values. Without loss of generality, we assume that each value composed of three integers: the first one between 0 and 2, the second one between 1 and β and the third one between 1 and 6. Informally, the property of the lemma will transfer from the origin variables $w_{i,1}, \ldots, w_{i,\beta}$ to the root variables $v_{i,1}, \ldots, v_{i,\beta}$. The transfer will be done line after line of the majority metaconstraint, similarly to the method used in the proof of the proposition. The first integer composing the value assigned to a majority variable corresponds to the trilean value to carry to the next line of the majority metaconstraint. The second integer composing the value assigned to a majority variable corresponds to the variable of the next line to carry the trilean value. The third integer composing the value assigned to a majority variable corresponds to one of the six possible bijections from $\{0, 1, 2\}$ to $\{0, 1, 2\}$. The correspondence is explicitely given in Table B.2.
   For example, if the third integer composing the value assigned to a majority variable is 3, then the bijection associated to this variable is f, with $f(0) = 1$, $f(1) = 0$ and $f(2) = 2$. So if the value assigned to the majority variable $m_{i,l,6}$ is composed of the three integers 1,8 and 3, in that order, then the trilean value 1 will be carried over to the majority variable $m_{i,l+1,8}$, and the first integer composing the value of the majority variable $m_{i,l+1,8}$ will be 0, because $f(1) = 0$, with f being the bijection associated to the majority variable $m_{i,l,6}$.

4. (guaranteeing permutations and bijections) For each $1 \leq l \leq \alpha + 1$, for all $1 \leq j \neq j' \leq \beta$, for each value a in the domain of the majority variable $m_{i,l,j}$ and each value b in the domain of the majority variable $m_{i,l,j'}$, a is incompatible with b if the second integer composing a is equal to the second integer composing b or if the third integer composing a is different from the third integer composing b.
   This is to ensure that in any majority constraint, the carrying of trilean values from line l to line $l+1$ follows a bijection, and that the bijections associated to the third integer in the values associated to the majority variables in line l of the metaconstraint are all the same. The latter property is needed to ensure that the property held by the first line of the majority metaconstraint, namely that one trilean value occurs at least $\alpha + 1$ times, is transferred to the last line of the majority metaconstraint, which is also the first line of the permutation constraints involving the variable v_i.
5. (first line in a majority metaconstraint) For each $1 \leq j \leq \beta$, the value 0 in the domain of the origin variable $w_{i,j}$ is incompatible with all values in the domain of the majority variable $m_{i,1,j}$ that do not have 0 as their first integer, the value 1 in the domain of the origin variable $w_{i,j}$ is incompatible with all values in the domain of the majority variable $m_{i,1,j}$ that do not have 1 as their first integer and the value 2 in the domain of the origin variable $w_{i,j}$ is incompatible with all values in the domain of the majority variable $m_{i,1,j}$ that do not have 2 as their first integer. This is to ensure that $m_{i,1,j}$ is the majority variable of the first line of the majority metaconstraint M_i that will carry the trilean value assigned to $w_{i,j}$ to the second line of the majority metaconstraint M_i.
6. (successive lines in a majority metaconstraint) For each $1 \leq l \leq \alpha$, for all $1 \leq j, j' \leq \beta$, let a be a value in the domain of the majority variable $m_{i,l,j}$ such that the second integer composing a is j', and let b be a value in the domain of the majority variable $m_{i,l+1,j'}$. Then a is incompatible with b if the first integer composing b is not the result of the bijection corresponding to the third integer composing a applied to the first integer composing a. The correspondence is given in Table B.2.
   For example, if the first integer composing a is 0, and the third integer composing a is 3, then a will be incompatible with b if the first integer composing b is 0 or 2, because the bijection corresponding to 3 (in the column labelled by '3') gives 1 when applied to 0 (in the line labelled by '0'). As another concrete example, if the first integer composing a is 2, and the third integer composing a is 6, then a will be incompatible with b if the first integer composing b is 1 or 2.
7. (last line in a majority metaconstraint) For all $1 \leq j, j' \leq \beta$, let a be a value in the domain of the majority variable $m_{i,\alpha+1,j}$ such that the second integer composing a is j', and let b be a value in the domain of the root variable $v_{i,j'}$. Then a is incompatible with b if b is not the result of the bijection corresponding to the third integer composing a applied to the first integer composing a. The correspondence is given in Table B.2.
8. (preventing contradictions around the origin variables) The majority variables on the first line in a majority constraint determine the values of the origin variables on the previous line, so we need to add the following constraints in order to make sure that we do not have a compatible assignment on two majority variables that implies a contradiction on one of the origin variables.
   For all $1 \leq j, j' \leq \alpha + 1$ such that $j \neq j'$, let a be a value in the domain of the origin variable $w_{i,j}$, let a' be a value in the domain of the majority variable $m_{i,1,j}$ and let b'

Table B.2: Correspondence between the third integer in the value assigned to majority variables and the bijections from $\{0, 1, 2\}$ to $\{0, 1, 2\}$.

trilean value \ third integer of a value assigned to a majority variable	1	2	3	4	5	6
0	0	0	1	1	2	2
1	1	2	0	2	0	1
2	2	1	2	0	1	0

be a value in the domain of the majority variable $m_{i,1,j'}$. If the first integer composing b' is not equal to a, then a and b' are incompatible. If the first integer composing a' is not equal to the first integer composing b', then a' and b' are incompatible.

9. (preventing contradictions around the root variables) The majority variables on the last line in a majority constraint determine the values of the root variables on the following line, so we need to add the following constraints in order to make sure that we do not have a compatible assignment on two majority or permutation variables that implies a contradiction on one of the root variables.
   9.1 For each $1 \leq j \leq n$ such that $j < i$, for all $1 \leq col, col' \leq \beta$, let a be a value in the domain of the permutation variable $c_{i,j,2\alpha,col}$ and let b be a value in the domain of the majority variable $m_{i,\alpha+1,col'}$ such that the second integer composing b is equal to $\lfloor a/3 \rfloor + 1$ (informally, both $c_{i,j,2\alpha,col}$ and $m_{i,\alpha+1,col'}$ carry a trilean value to the same root variable $v_{i,\lfloor a/3 \rfloor+1}$). Then a and b are incompatible if the result of the bijection corresponding to the third integer composing b applied to the first integer composing b is not congruent to a modulo 3. The correspondence is given in Table B.2.
   9.2 For each $1 \leq j \leq n$ such that $i < j$, for all $1 \leq col, col' \leq \beta$, let a be a value in the domain of the majority variable $m_{i,\alpha+1,col}$ such that the second integer composing a is equal to col' and let b be a value in the domain of the permutation variable $c_{i,j,1,col'}$. Then a and b are incompatible if the result of the bijection corresponding to the third integer composing a applied to the first integer composing a is not congruent to b modulo 3.
10. (the rest) All couples of values containing at least one value in M_i that have not had their compatibility specified yet, including the couples that contain a value from another majority metaconstraint or from a permutation metaconstraint, are set to compatible.

To illustrate the construction, we present in Figure B.4 the majority metaconstraint M_i. This figure should be viewed as a part of the same instance as the one from Figure 3, and i takes the same value. In particular, the root variables $v_{i,1}, \ldots, v_{i,\beta}$ appear in both figures. We present an example of a possible assignment for the 63 variables composing the majority metaconstraint, with the values assigned to the 9 root variables being the same as they were in Figure 3. Note that in this gadget, any compatible assignment on $\alpha = 4$ variables can be extended to a partial solution on all 63 variables.

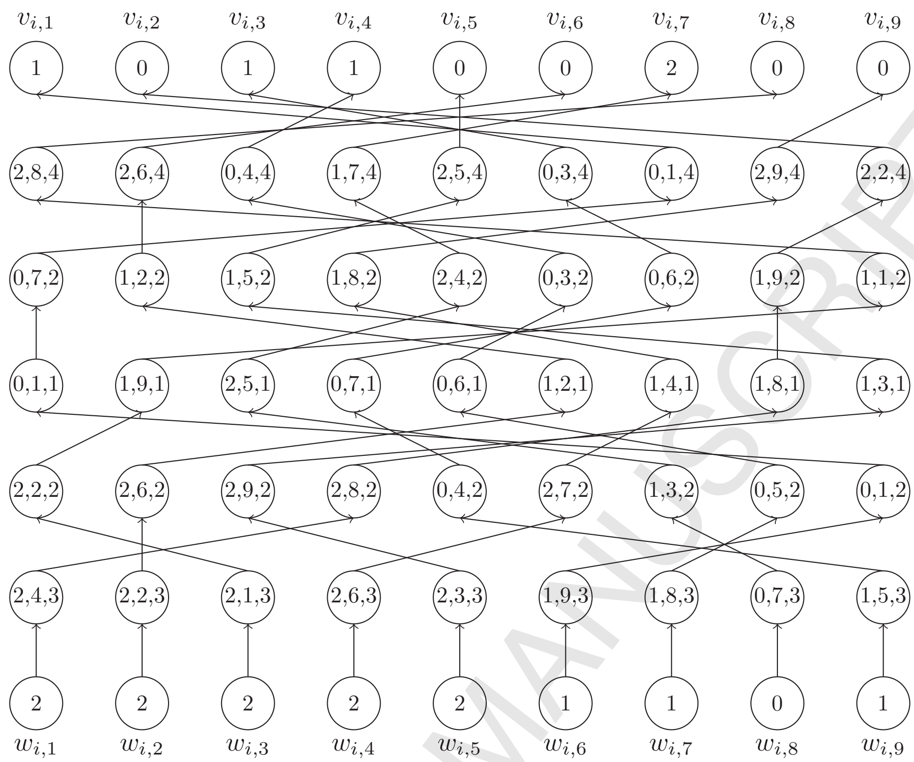

Figure B.4: The majority metaconstraint M_i.

It is clear from the definition of M_i (in particular from the fourth point) that if the bottom line fulfills the absolute majority property, namely that one value is assigned at least $\alpha + 1$ times among the $2\alpha + 1$ assignments of the line, then every subsequent line until the top one will also fulfill that property. The proof that adding the n majority metaconstraints conserves α^- minimality can be done by keeping the exact same construction and lemmas used in Appendix A.